PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/00 — A1

(11) International Publication Number: WO 00/27861

(43) International Publication Date: 18 May 2000 (18.05.00)

(21) International Application Number: PCT/US99/26860

(22) International Filing Date: 12 November 1999 (12.11.99)

(30) Priority Data:
60/108,255 12 November 1998 (12.11.98) US

(71) Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY [US/US]; Suite 350, 900 Welch Road, Palo Alto, CA 94304 (US).

(72) Inventors: CONTI, Marco; 24 Ryan Court, Stanford, CA 94305 (US). PAHLKE, Gudrun; Apartment #10, 806 Coleman Avenue, Menlo Park, CA 94025 (US).

(74) Agent: FIELD, Bret, E.; Bozicevic, Field & Francis LLP, Suite 200, 285 Hamilton Avenue, Palo Alto, CA 94301 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: NOVEL PHOSPHODIESTERASE INTERACTING PROTEINS

(57) Abstract

Nucleic acid compositions encoding novel PDE interacting proteins, as well as the novel PDE interacting proteins themselves, are provided. Also provided are methods of producing the subject nucleic acid and protein compositions. The subject polypeptide and nucleic acid compositions find use in a variety of applications, including research, diagnostic, and therapeutic agent screening applications, as well as in treatment therapies for disease conditions associated with PDE activity, particularly inflammatory diseases.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

NOVEL PHOSPHODIESTERASE INTERACTING PROTEINS

ACKNOWLEDGMENT OF GOVERNMENT SUPPORT

This invention was made with Government support under Grant No. HD20788 awarded by the National Institutes of Health. The Government has certain rights in this invention.

INTRODUCTION

Field of the Invention

The field of the invention is cyclic nucleotide phosphodiesterases, particularly cAMP phosphodiesterases.

Background of the Invention

Cyclic nucleotide phosphodiesterases are a class of enzymes that catalyze the hydrolysis of phosphodiester bonds in cyclic nucleotides, e.g. cAMP. Cyclic nucleotides are important second messengers that regulate and mediate a number of cellular responses to extracellular signals, such as hormones, light and neurotransmitters. Since cyclic nucleotide phosphodiesterases modulate the concentration of cyclic nucleotides, these enzymes play a significant role in signal transduction. There are at least ten different classes of cyclic phosphodiesterases, seven of which are: (I) Ca(2+)/calmodulin-dependent PDEs; (II) cGMP-stimulated PDEs; (III) cGMP-inhibited PDEs; (IV) cAMP-specific PDEs; (V) cGMP-specific PDEs; (VI) photoreceptor PDEs; and (VII) high-affinity, cAMP-specific PDEs. Because of their role in signal transduction, cyclic nucleotide phosphodiesterases have been pursued as therapeutic or pharmacologic targets in the modulation of a variety of distinct physiological processes.

cAMP phosphodiesterase inhibitors hold great promise as therapeutic agents for use in the treatment of inflammation. Specifically, data indicates that these types of inhibitors are as effective, or even more effective, than adrenal steroids in suppressing most functions of inflammatory cells, including: migration, adhesion and secretion of cytokines. Specific cAMP phosphodiesterase inhibitors that have been studied include: rolipram, theophylline, and the like. In addition, research is ongoing to identify new cAMP phosphodiesterase inhibitors.

Despite their promise as anti-inflammatory therapeutic agents, cAMP-phosphodiesterase inhibitors identified to date have demonstrated significant toxic side effects that have limited to their generalized use in the treatment of inflammation.

As such, there is continued interest in the identification of new, more selective cAMP phosphodiesterase inhibitors for potential use as anti-inflammatory therapeutic agents. These efforts have employed recombinant phosphodiesterases for automated screening of candidate agents. Use of recombinant phosphodiesterases in screening applications has, however, been problematic as such recombinant enzymes have altered conformation as compared to their naturally occurring counterparts, which affects the interaction with potential inhibitors and thereby confounds the results that are obtained. As such, the screening results obtained by using such recombinant proteins are problematic.

Therefore, there is much interest in the further elucidation of the conformation of phosphodiesterases and other factors that may modulate the interaction of these enzymes with inhibitors.

Relevant Literature

The role of cAMP phosphodiesterases in inflammatory processes is reviewed in Torphy, Am. J. Respir. Crit. Care Med. (1998) 157:351-370. *See also* Houslay et al., Adv. Pharmacol (1998) 44: 225-342 and Spina et al., Adv. Pharmacol (1998) 44: 33-89, as well as U.S. Patent No. 5,798,373, the disclosure of which is herein incorporated by reference.

SUMMARY OF THE INVENTION

Nucleic acid compositions encoding phosphodiesterase interacting proteins, e.g. myomegalin, as well as the polypeptide compositions encoded thereby, are provided. Also provided are complexes of the subject phosphodiesterase interacting protein with a phosphodiesterase enzyme. The subject polypeptide and nucleic acid compositions, as well as complexes thereof, find use in a variety of applications, including research, diagnostic, and therapeutic agent identification and screening applications, as well as in therapeutic applications.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 provides the amino acid sequence of rat myomegalin.

Figure 2 provides the cDNA sequence of a clone having an open reading frame encoding the myomegalin protein having the amino acid sequence of Figure 1.

Figure 3 provides the nucleic acid sequence from the first met to the first stop codon in the sequence of Figure 2.

Figure 4 provides the nucleic acid sequence of human myomegalin.

Figure 5 provides the amino acid sequence of human myomegalin.

Figure 6 provides the amino acid sequence of rat M14 protein.

DETAILED DESCRIPTION OF THE INVENTION

Novel phosphodiesterase interacting proteins, particularly myomegalin, as well as nucleic acid compositions encoding the same, are provided. Also provided are complexes of the subject proteins and phosphodiesterases. The subject polypeptide and nucleic acid compositions find use in a variety of applications, including research, diagnostic, and therapeutic agent identification and screening applications, as well as in therapeutic applications.

Before the subject invention is described further, it is to be understood that the invention is not limited to the particular embodiments of the invention described below, as variations of the particular embodiments may be made and still fall within the scope of the appended claims. It is also to be understood that the terminology employed is for the purpose of describing particular embodiments, and is not intended to be limiting. Instead, the scope of the present invention will be established by the appended claims.

In this specification and the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs.

NUCLEIC ACID COMPOSITIONS

Nucleic acid compositions encoding phosphodiesterase (PDE) interacting proteins, as well as fragments thereof, are provided. The subject nucleic acid compositions encode proteins that interact with a phoshodiesterase enzyme, modulate its conformation and direct

its location in a cell. In other words, the proteins encoded by the subject nucleic acid compositions are those that target a (PDE) to a particular subcellular compartment and alter the function and/or properties of the PDE. Of particular interest are nucleic acid compositions which encode proteins that bind to a PDE IV isoenzyme, including PDE4A, PDE4B, PDE4C, PDE4D, and the like.

By nucleic acid composition is meant a composition comprising a sequence of DNA having an open reading frame that encodes a PDE interacting polypeptide, i.e. a gene encoding a polypeptide that interacts with a PDE (e.g. binds to and targets a PDE), and is capable, under appropriate conditions, of being expressed as a PDE interacting polypeptide. Also encompassed in this term are nucleic acids that are homologous, substantially similar or identical to the nucleic acids encoding PDE interacting polypeptides or proteins. Thus, the subject invention provides genes encoding mammalian PDE interacting proteins, such as genes encoding human PDE interacting polypeptides and homologs thereof, as well as non-human mammalian PDE interacting polypeptides and homologs thereof, e.g. rat and mouse proteins.

Of particular interest is a nucleic acid composition encoding a myomegalin protein, particularly a mammalian myomegalin protein, described in greater detail *infra*, or a fragment or homolog thereof. Specific nucleic acid compositions of interest include: polynucleotides encoding a rat myomegalin protein, such as polynucleotides having a nucleotide sequence found in SEQ ID NOs: 1 or 3, including polynucleotides in which the entire sequence is the same as the sequence of SEQ ID NOs. 1 or 3; and polynucleotides encoding human myomegalin protein, such as polynucleotides having a nucleotide sequence found in SEQ ID NO:04, including polynucleotides in which the entire sequence is the same as the sequence of SEQ ID NOs. 04, as well as those in which the entire sequence is the same as the sequence of an ORF found in SEQ ID NO:04.

Also of interest are nucleic acid compositions encoding an M14 polypeptide, described in greater detail *infra*, or a fragment or homolog thereof. Specific nucleic acid compositions of interest include polynucleotides encoding a rat M14 polypeptide, such as polynucleotides encoding an M14 polypeptide having the amino acid sequence set forth in SEQ ID NO:08. Polynucleotides encoding M14 homologs, and polynucleotides encoding PDE-interacting fragments of an M14 polypeptide, are also of interest.

Also of interest are nucleic acid compositions encoding a huntingtin-interacting protein, e.g., HIP1. Specific nucleic acid compositions of interest include a polynucleotide encoding a human HIP1 polypeptide, including, for example, a polynucleotide as disclosed in GenBank Accession No. U79734.

The source of homologous genes to those specifically listed above may be any mammalian species, e.g., primate species, particularly human; rodents, such as guinea pigs and mice, canines, felines, bovines, ovines, equines, yeast, nematodes, etc. Between mammalian species, *e.g.*, human and mouse, homologs have substantial sequence similarity, *e.g.* at least 75% sequence identity, usually at least 90%, more usually at least 95% between nucleotide sequences. Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, *etc.* A reference sequence will usually be at least about 18 nt long, more usually at least about 30 nt long, and may extend to the complete sequence that is being compared. Algorithms for sequence analysis are known in the art, such as BLAST, described in Altschul *et al.* (1990), *J. Mol. Biol.* **215**:403-10. Unless stated otherwise herein, all sequence identity figures provided in this application are determined using the BLAST program at default settings (e.g. w=4; T=17). The sequences provided herein are essential for recognizing genes encoding PDE interacting protein-related and homologous polynucleotides in database searches.

Nucleic acids encoding the subject PDE interacting proteins and polypeptides of the subject invention may be cDNAs or genomic DNAs, as well as fragments thereof. Also provided are genes comprising the subject nucleic acid compositions, where the term "gene" shall be intended to mean the open reading frame encoding specific PDE interacting proteins and polypeptides, and introns, as well as adjacent 5′ and 3′ non-coding nucleotide sequences involved in the regulation of expression, up to about 20 kb beyond the coding region, but possibly further in either direction. The gene may be introduced into an appropriate vector for extrachromosomal maintenance or for integration into a host genome.

The term "cDNA" as used herein is intended to include all nucleic acids that share the arrangement of sequence elements found in native mature mRNA species, where sequence elements are exons and 3′ and 5′ non-coding regions. Normally mRNA species have contiguous exons, with the intervening introns, when present, being removed by nuclear RNA splicing, to create a continuous open reading frame encoding an PDE interacting protein.

A genomic sequence of interest comprises the nucleic acid present between the initiation codon and the stop codon, as defined in the listed sequences, including all of the introns that are normally present in a native chromosome. It may further include the 3′ and 5′ untranslated regions found in the mature mRNA. It may further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, *etc.*, including about 1 kb, but possibly more, of flanking genomic DNA at either the 5′ or 3′ end of the transcribed region. The genomic DNA may be isolated as a fragment of 100 kbp or smaller; and substantially free of flanking chromosomal sequence. The genomic DNA flanking the coding region, either 3' or 5', or internal regulatory sequences as sometimes found in introns, contains sequences required for proper tissue and stage specific expression.

The nucleic acid compositions of the subject invention may encode all or a part of the subject PDE interacting proteins and polypeptides, described in greater detail *infra*. Double or single stranded fragments may be obtained from the DNA sequence by chemically synthesizing oligonucleotides in accordance with conventional methods, by restriction enzyme digestion, by PCR amplification, *etc.* For the most part, DNA fragments will be of at least 15 nt, usually at least 18 nt or 25 nt, and may be at least about 50 nt.

The genes of the subject invention are isolated and obtained in substantial purity, generally as other than an intact chromosome. Usually, the DNA will be obtained substantially free of other nucleic acid sequences that do not include a sequence encoding a PDE interacting protein or fragment thereof, generally being at least about 50%, usually at least about 90% pure and are typically "recombinant," *i.e.* flanked by one or more nucleotides with which it is not normally associated on a naturally occurring chromosome.

In addition to the plurality of uses described in greater detail in following sections, the subject nucleic acid compositions find use in the preparation of all or a portion of the PDE interacting polypeptides, as described below.

POLYPEPTIDE COMPOSITIONS

Also provided by the subject invention are PDE interacting proteins and polypeptides, i.e. proteins and polypeptides that are capable of binding to and modulating PDEs, specifically cAMP-PDEs, and more particularly cAMP-PDE4 isoforms, such as PDE4A, PDE4B, PDE4C, PDE4D, and the like.

The term polyeptide composition as used herein refers to both the full length proteins as well as portions or fragments thereof. Also included in this term are variations of the naturally occurring proteins, where such variations are homologous or substantially similar to the naturally occurring protein, as described in greater detail below, be the naturally occurring protein the human protein, rat protein, or protein from some other species which naturally expresses an PDE interacting protein, usually a mammalian species. In the following description of the subject invention, the term PDE interacting protein is used to refer not only to the human form of such proteins, but also to homologs thereof expressed in non-human species, e.g. murine, rat and other mammalian species.

The subject PDE proteins are, in their natural environment, capable of modulating the form/function of PDEs, as well as targeting PDEs to specific subcellular compartments within a cell. In many embodiments, the subject PDE interacting proteins serve as PDE anchoring proteins.

In many embodiments, the subject proteins are characterized by the presence of one or more coiled domains and leucine zippers. Furthermore, in certain embodiments, e.g. certain rat myomegalin proteins, the subject proteins have a region of high homology with *Drosophila* centrosomin, whereby high homology is meant at least about 30, usually at least about 40 % sequence identity.

In many embodiments, the proteins range in length from about 1500 to 3000, usually from about 1600 to 2800 and more usually from about 1650 to 2600 amino acid residues, and the projected molecular weight of the subject proteins based solely on the number of amino acid residues in the protein ranges from about 150 to 320, usually from about 160 to 300 kDa, where the actual molecular weight may vary depending on the amount of glycolsylation, if any, of the protein and the apparent molecular weight may be considerably less (40 to 50 kDa) due to SDS binding on gels. On other embodiments, the length of the proteins may be much smaller, e.g. as in the case of splice variants or post translated products, where the length in these proteins may be as short as 40%, usually no shorter than about 50% of the above lengths.

Of particular interest in many embodiments are proteins that are non-naturally glycosylated. By non-naturally glycosylated is meant that the protein has a glycosylation pattern, if present, which is not the same as the glycosylation pattern found in the corresponding naturally occurring protein. For example, a human phosphodiesterase binding

protein of the subject invention and of this particular embodiment is characterized by having a glycosylation pattern, if it is glycosylated at all, that differs from that of naturally occurring human PDE binding protein. Thus, the non-naturally glycosylated PDE interacting or binding proteins of this embodiment include non-glycosylated PDE interacting proteins, i.e. proteins having no covalently bound glycosyl groups.

A PDE interacting protein of the subject invention of particular interest is myomegalin, particularly mammalian myomegalin and more particuarly, rat or human myomegalin. In many embodiments, mammalian myomegalin ranges in length from about 2000 to 3000, usually from about 2200 to 2800 and more usually from about 2300 to 2600 aa residues. The projected molecular weight of these myomegalin proteins based solely on the number of amino acid residues in the protein ranges from about 220 to 320, usually from about 220 to 300 and more usually from about 240 to 300 kDa, where the actual molecular weight may vary depending on the amount of glycolsylation, if any, of the protein and the apparent molecular weight may be considerably less (40 to 50 kDa) due to SDS binding on gels. Also of interest are mammalian myomegalin proteins that are shorter than those described above, where these shorter proteins could be splice variants or the products of post-translational activity, and the like.

Of particular interest in certain embodiments is the rat myomegalin protein, where the rat myomegalin protein of the subject invention has an amino acid sequence that is substantially the same as or identical to the sequence appearing as SEQ ID NO:02 *infra* and appearing in Figure 1. By substantially the same as is meant a protein having a sequence that has at least about 80%, usually at least about 90% and more usually at least about 98% sequence identity with the sequence of SED ID NO:02. Also of particular interest is an approximately 65 kDa rat myomegalin protein expressed in rat testis. Yet another protein of particular interest is the human myomegalin protein of the subject invention which has an amino acid sequence that is substantially the same as or identical to the sequence appearing as SEQ ID NO:05 *infra* and appearing in Figure 5. By substantially the same as is meant a protein having a sequence that has at least about 80%, usually at least about 90% and more usually at least about 98% sequence identity with the sequence of SED ID NO:05.

Another PDE interacting protein of the subject invention of particular interest is M14, particularly mammalian M14, and more particularly, rat or human M14. In many embodiments, mammalian M14 ranges in length from about 1500 to about 2000, usually from

about 1600 to about 1800, usually from about 1650 to about 1700, and more usually from about 1670 to about 1690 amino acid residues. The projected molecular weight of these M14 polypeptides, based solely on the number of amino acid residues in the protein, ranges from about 150 to about 200 kDa, usually from about 160 to about 180 kDa, usually from about 165 to about 170 kDa. Rat M14 protein has a mobility on SDS-PAGE of about 185 kDa. The actual molecular weight may vary depending on the amount of glycosylation or other post-translational modifications, if any, of the protein, and the apparent molecular weight may be considerably less (e.g. 40-50 kDa) due to SDS binding on gels. Also of interest are PDE-interacting fragments of the above-described M14 proteins.

Of particular interest in certain embodiments is a rat M14 protein, where the rat M14 protein of the subject invention has an amino acid sequence that is substantially the same or identical to the sequence set forth in SEQ ID NO:08 and appearing in Figure 6. By substantially the same as is meant a protein having a sequence that has at least about 80%, usually at least about 90% and more usually at least about 98% sequence identity with the sequence of SED ID NO:08. Proteins homologous to rat M14 are also of interest, including, e.g., an Ese2L protein as described in Sengar et al. (1999) *EMBO J.* 18:1159-1171.

Also of interest are huntingtin interacting proteins, and PDE-interacting fragments, variants and homologs thereof. In some embodiments, huntingtin interacting protein (HIP) is a human HIP1 protein having an amino acid sequence as disclosed in GenBank Accession No. U79734, The human HIP1 protein is described in Kalchman et al. (1997) *Nature Genetics* 16:44-53.

In addition to the specific PDE interacting proteins described above, homologs or proteins (or fragments thereof) from other species, i.e. other animal or plant species, are also provided, where such homologs or proteins may be from a variety of different types of species, usually mammals, e.g. rodents, such as mice, rats; domestic animals, e.g. horse, cow, dog, cat; and humans. By homolog is meant a protein having at least about 35 %, usually at least about 40% and more usually at least about 60 % amino acid sequence identity with a specific PDE interacting protein as identified in: (a) SEQ ID NO: 02 and appearing in Figure 1; or (b) SEQ ID NO:05 and appearing in Figure 5; or (c) SEQ ID NO:08 and appearing in Figure 6.

The PDE interacting proteins of the subject invention (e.g. human myomegalin, rat myomegalin or homologs thereof) are present in a non-naturally occurring environment, *e.g.*

are separated from their naturally occurring environment. In certain embodiments, the subject protein is present in a composition that is enriched for the subject protein as compared to the protein in its naturally occurring environment. As such, purified PDE interacting protein is provided, where by purified is meant that PDE interacting protein is present in a composition that is substantially free of non PDE interacting proteins, where by substantially free is meant that less than 90 %, usually less than 60 % and more usually less than 50 % of the composition is made up of non-PDE interacting proteins.

In certain embodiments of interest, the PDE interacting protein is present in a composition that is substantially free of the constituents that are present in its naturally occurring environment. For example, a human PDE interacting protein comprising composition according to the subject invention in this embodiment will be substantially, if not completely, free of those other biological constituents, such as proteins, carbohydrates, lipids, etc., with which it is present in its natural environment. As such, protein compositions of these embodiments will necessarily differ from those that are prepared by purifying the protein from a naturally occurring source, where at least trace amounts of the protein's constituents will still be present in the composition prepared from the naturally occurring source.

The PDE interacting protein of the subject invention may also be present as an isolate, by which is meant that the PDE interacting protein is substantially free of both non-PDE interacting proteins and other naturally occurring biologic molecules, such as oligosaccharides, polynucleotides and fragments thereof, and the like, where substantially free in this instance means that less than 70 %, usually less than 60% and more usually less than 50 % of the composition containing the isolated PDE interacting protein is a non-PDE interacting protein naturally occurring biological molecule. In certain embodiments, the subject protein is present in substantially pure form, where by substantially pure form is meant at least 95%, usually at least 97% and more usually at least 99% pure.

In addition to the naturally occurring proteins, polypeptides which vary from the naturally occurring proteins are also provided. By polypeptides is meant proteins having an amino acid sequence encoded by an open reading frame (ORF) of an gene according to the subject invention, described *supra*, including the full length protein and fragments thereof, particularly biologically active fragments and/or fragments corresponding to functional domains; and including fusions of the subject polypeptides to other proteins or parts thereof. Fragments of interest will typically be at least about 10 aa in length, usually at least about 50

aa in length, and may be as long as 300 aa in length or longer, but will usually not exceed about 1000 aa in length, where the fragment will have a stretch of amino acids that is identical to the protein of SEQ ID NO:02, SEQ ID NO:05, or SEQ ID NO:08, or a homolog thereof, of at least about 10 aa, and usually at least about 15 aa, and in many embodiments at least about 50 aa in length.

PREPARATION OF PDE INTERACTING POLYPEPTIDES

The subject PDE interacting proteins and polypeptides may be obtained from naturally occurring sources or synthetically produced. Where obtained from naturally occurring sources, the source chosen will generally depend on the species from which the PDE interacting protein is to be derived, e.g. muscle tissue, heart tissue, brain tissue, testis tissue, and the like.

The subject PDE interacting polypeptide compositions may be synthetically derived by expressing a recombinant gene encoding the PDE interacting protein, such as the polynucleotide compositions described above, in a suitable host. For expression, an expression cassette may be employed. The expression vector will provide a transcriptional and translational initiation region, which may be inducible or constitutive, where the coding region is operably linked under the transcriptional control of the transcriptional initiation region, and a transcriptional and translational termination region. These control regions may be native to the gene encoding the particular PDE interacting protein, or may be derived from exogenous sources.

Expression vectors generally have convenient restriction sites located near the promoter sequence to provide for the insertion of nucleic acid sequences encoding heterologous proteins. A selectable marker operative in the expression host may be present. Expression vectors may be used for the production of fusion proteins, where the exogenous fusion peptide provides additional functionality, i.e. increased protein synthesis, stability, reactivity with defined antisera, an enzyme marker, e.g. β-galactosidase, etc.

Expression cassettes may be prepared comprising a transcription initiation region, the gene or fragment thereof, and a transcriptional termination region. Of particular interest is the use of sequences that allow for the expression of functional epitopes or domains, usually at least about 8 amino acids in length, more usually at least about 15 amino acids in length, to about 25 amino acids, and up to the complete open reading frame of the gene. After

introduction of the DNA, the cells containing the construct may be selected by means of a selectable marker, the cells expanded and then used for expression.

The subject proteins and polypeptides may be expressed in prokaryotes or eukaryotes in accordance with conventional ways, depending upon the purpose for expression. For large scale production of the protein, a unicellular organism, such as *E. coli, B. subtilis, S. cerevisiae*, insect cells in combination with baculovirus vectors, or cells of a higher organism such as vertebrates, particularly mammals, *e.g.* COS 7 cells, may be used as the expression host cells. In some situations, it is desirable to express the subject proteins in eukaryotic cells, where the protein will benefit from native folding and post-translational modifications. Small peptides can also be synthesized in the laboratory. Polypeptides that are subsets of the complete protein sequence may be used to identify and investigate parts of the protein important for function.

Once the source of the protein is identified and/or prepared, e.g. a transfected host expressing the protein is prepared, the protein is then purified to produce the desired PDE interacting protein comprising composition. Any convenient protein purification procedures may be employed, where suitable protein purification methodologies are described in Guide to Protein Purification, (Deuthser ed.) (Academic Press, 1990). For example, a lysate may be prepared from the original source, e.g. naturally occurring cells or tissues that express a PDE interacting protein or the expression host expressing the PDE interacting protein, and purified using HPLC, exclusion chromatography, gel electrophoresis, affinity chromatography, and the like.

USES OF THE SUBJECT POLYPEPTIDE AND NUCLEIC ACID COMPOSITIONS

The subject polypeptide and nucleic acid compositions find use in a variety of different applications, including diagnostic, and therapeutic agent screening/discovery/preparation applications, as well as the treatment of disease conditions associated with PDE interacting protein activity.

GENERAL APPLICATIONS

The subject nucleic acid compositions find use in a variety of applications, including: (a) the identification of PDE interacting protein gene homologs, e.g. myomegalin homologs; (b) as a source of novel promoter elements; (c) the identification of PDE interacting protein

expression regulatory factors; (d) as probes and primers in hybridization applications, e.g. PCR; (e) the identification of expression patterns in biological specimens; (f) the preparation of cell or animal models for PDE interacting protein function; (g) the preparation of *in vitro* models for PDE interacting protein function; etc.

Identification of homologs

Homologs of the PDE interacting protein gene, e.g. the myomegalin gene, or the M14 gene, are identified by any of a number of methods. A fragment of the provided cDNA may be used as a hybridization probe against a cDNA library from the target organism of interest, where low stringency conditions are used. The probe may be a large fragment, or one or more short degenerate primers. Nucleic acids having sequence similarity are detected by hybridization under low stringency conditions, for example, at 50°C and 6×SSC (0.9 M sodium chloride/0.09 M sodium citrate) and remain bound when subjected to washing at 55°C in 1×SSC (0.15 M sodium chloride/0.015 M sodium citrate). Sequence identity may be determined by hybridization under stringent conditions, for example, at 50°C or higher and 0.1×SSC (15 mM sodium chloride/01.5 mM sodium citrate). Nucleic acids having a region of substantial identity to the provided sequences, e.g. allelic variants, genetically altered versions of the gene, etc., bind to the provided sequences under stringent hybridization conditions. By using probes, particularly labeled probes of DNA sequences, one can isolate homologous or related genes.

Identification of Novel Promoter Elements

The sequence of the 5' flanking region may be utilized for promoter elements, including enhancer binding sites, that provide for regulation in tissues where the subject gene is expressed. The tissue specific expression is useful for determining the pattern of expression, and for providing promoters that mimic the native pattern of expression. Naturally occurring polymorphisms in the promoter region are useful for determining natural variations in expression, particularly those that may be associated with disease.

Identification of Expression Regulatory Factors

Alternatively, mutations may be introduced into the promoter region to determine the effect of altering expression in experimentally defined systems. Methods for the identification

of specific DNA motifs involved in the binding of transcriptional factors are known in the art, *e.g.* sequence similarity to known binding motifs, gel retardation studies, *etc.* For examples, see Blackwell *et al.* (1995), *Mol. Med.* 1:194-205; Mortlock *et al.* (1996), *Genome Res.* 6:327-33; and Joulin and Richard-Foy (1995), *Eur. J. Biochem.* 232:620-626.

The regulatory sequences may be used to identify *cis* acting sequences required for transcriptional or translational regulation of expression of the subject gene, e.g. the myomegalin gene, especially in different tissues or stages of development, and to identify *cis* acting sequences and *trans*-acting factors that regulate or mediate expression of the subject gene. Such transcription or translational control regions may be operably linked to a gene of the subject invention in order to promote expression of wild type or altered PDE interacting protein, e.g. myomegalin, or other proteins of interest in cultured cells, or in embryonic, fetal or adult tissues, and for gene therapy.

Probes and Primers

Small DNA fragments are useful as primers for PCR, hybridization screening probes, *etc.* Larger DNA fragments, *i.e.* greater than 100 nt are useful for production of the encoded polypeptide, as described in the previous section. For use in amplification reactions, such as PCR, a pair of primers will be used. The exact composition of the primer sequences is not critical to the invention, but for most applications the primers will hybridize to the subject sequence under stringent conditions, as known in the art. It is preferable to choose a pair of primers that will generate an amplification product of at least about 50 nt, preferably at least about 100 nt. Algorithms for the selection of primer sequences are generally known and are available in commercial software packages. Amplification primers hybridize to complementary strands of DNA, and will prime towards each other.

Identification of Expression Patterns in Biological Specimens

The DNA may also be used to identify expression of the gene in a biological specimen. The manner in which one probes cells for the presence of particular nucleotide sequences, as genomic DNA or RNA, is well established in the literature. Briefly, DNA or mRNA is isolated from a cell sample. The mRNA may be amplified by RT-PCR, using reverse transcriptase to form a complementary DNA strand, followed by polymerase chain reaction amplification using primers specific for the subject DNA sequences. Alternatively, the mRNA

sample is separated by gel electrophoresis, transferred to a suitable support, *e.g.* nitrocellulose, nylon, *etc.*, and then probed with a fragment of the subject DNA as a probe. Other techniques, such as oligonucleotide ligation assays, *in situ* hybridizations, and hybridization to DNA probes arrayed on a solid chip may also find use. Detection of mRNA hybridizing to the subject sequence is indicative of gene expression in the sample.

The Preparation of PDE Interacting Protein Mutants

The sequence of a gene according to the subject invention, including flanking promoter regions and coding regions, may be mutated in various ways known in the art to generate targeted changes in promoter strength, sequence of the encoded protein, *etc.* The DNA sequence or protein product of such a mutation will usually be substantially similar to the sequences provided herein, *i.e.* will differ by at least one nucleotide or amino acid, respectively, and may differ by at least two but not more than about ten nucleotides or amino acids. The sequence changes may be substitutions, insertions, deletions, or a combination thereof. Deletions may further include larger changes, such as deletions of a domain or exon. Other modifications of interest include epitope tagging, *e.g.* with the FLAG system, HA, *etc.* For studies of subcellular localization, fusion proteins with green fluorescent proteins (GFP) may be used.

Techniques for *in vitro* mutagenesis of cloned genes are known. Examples of protocols for site specific mutagenesis may be found in Gustin *et al. (1993)*, *Biotechniques* **14**:22; Barany (1985), *Gene* **37**:111-23; Colicelli *et al. (1985)*, *Mol. Gen. Genet.* **199**:537-9; and Prentki *et al.* (1984), *Gene* **29**:303-13. Methods for site specific mutagenesis can be found in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, CSH Press 1989, pp. 15.3-15.108; Weiner *et al.* (1993), *Gene* **126**:35-41; Sayers *et al.* (1992), *Biotechniques* **13**:592-6; Jones and Winistorfer (1992), *Biotechniques* **12**:528-30; Barton *et al.* (1990), *Nucleic Acids Res* **18**:7349-55; Marotti and Tomich (1989), *Gene Anal. Tech.* **6**:67-70; and Zhu (1989), *Anal Biochem* **177**:120-4. Such mutated genes may be used to study structure-function relationships of PDE interacting proteins, or to alter properties of the protein that affect its function or regulation.

Production of *In Vivo* Models of PDE Interacting Protein Function

The subject nucleic acids can be used to generate transgenic, non-human animals or site specific gene modifications in cell lines. Transgenic animals may be made through homologous recombination, where the normal PDE interacting protein gene locus is altered. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACs, and the like.

The modified cells or animals are useful in the study of PDE interacting protein function and regulation. For example, a series of small deletions and/or substitutions may be made in the host's native PDE interacting protein gene to determine the role of different exons in cholesterol metabolism, e.g. cholesterol ester synthesis, cholesterol absorption, *etc.* Specific constructs of interest include anti-sense constructs which will block PDE interacting protein expression, expression of dominant negative gene mutations, and over-expression of PDE interacting protein genes. Where a particular genetic sequence is introduced, the introduced sequence may be either a complete or partial sequence of an PDE interacting protein gene native to the host, or may be a complete or partial sequence that is exogenous to the host animal, *e.g.*, a human sequence. A detectable marker, such as *lac Z*, may be introduced into the locus, where upregulation of gene expression will result in an easily detected change in phenotype.

One may also provide for expression of the gene or variants thereof in cells or tissues where it is not normally expressed, at levels not normally present in such cells or tissues, or at abnormal times of development.

DNA constructs for homologous recombination will comprise at least a portion of the gene native to the species of the host animal, wherein the gene has the desired genetic modification(s), and includes regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For various techniques for transfecting mammalian cells, see Keown *et al.* (1990), *Meth. Enzymol.* **185**:527-537.

For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, *e.g.* mouse, rat, guinea pig, *etc.* Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of leukemia inhibiting factor

(LIF). When ES or embryonic cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting offspring screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected.

The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogeneic or congenic grafts or transplants, or in *in vitro* culture. The transgenic animals may be any non-human mammal, such as laboratory animals, domestic animals, etc. The transgenic animals may be used in functional studies, drug screening, *etc.*, *e.g.* to determine the effect of a candidate drug on PDE interacting binding protein activity and/or the enzymatic activity of the PDE/PDE interacting protein complex.

Production of *In Vitro* Models of PDE Interacting Protein Function

One can also use the polypeptide compositions of the subject invention to produce *in vitro* models of PDE interacting protein function. In addition to the subject PDE interacting protein, such models will generally include at least a PDE as well as a cyclic nucleotide, and a means to monitor the activity of the enzyme in the presence of the PDE interacting protein, e.g. a labeled isotope, etc.

DIAGNOSTIC APPLICATIONS

Also provided are methods of diagnosing disease states associated with PDE interacting protein activity, e.g. based on observed levels of PDE interacting protein or the expression level of the gene in a biological sample of interest. Samples, as used herein, include

biological fluids such as semen, blood, cerebrospinal fluid, tears, saliva, lymph, dialysis fluid and the like; organ or tissue culture derived fluids; and fluids extracted from physiological tissues. Also included in the term are derivatives and fractions of such fluids. The cells may be dissociated, in the case of solid tissues, or tissue sections may be analyzed. Alternatively a lysate of the cells may be prepared.

A number of methods are available for determining the expression level of a gene or protein in a particular sample. Diagnosis may be performed by a number of methods to determine the absence or presence or altered amounts of normal or abnormal PDE interacting protein in a patient sample. For example, detection may utilize staining of cells or histological sections with labeled antibodies, performed in accordance with conventional methods. Cells are permeabilized to stain cytoplasmic molecules. The antibodies of interest are added to the cell sample, and incubated for a period of time sufficient to allow binding to the epitope, usually at least about 10 minutes. The antibody may be labeled with radioisotopes, enzymes, fluorescers, chemiluminescers, or other labels for direct detection. Alternatively, a second stage antibody or reagent is used to amplify the signal. Such reagents are well known in the art. For example, the primary antibody may be conjugated to biotin, with horseradish peroxidase-conjugated avidin added as a second stage reagent. Alternatively, the secondary antibody conjugated to a flourescent compound, *e.g.* fluorescein, rhodamine, Texas red, *etc.* Final detection uses a substrate that undergoes a color change in the presence of the peroxidase. The absence or presence of antibody binding may be determined by various methods, including flow cytometry of dissociated cells, microscopy, radiography, scintillation counting, *etc.*

Alternatively, one may focus on the expression of the gene. Biochemical studies may be performed to determine whether a sequence polymorphism in an coding region or control regions is associated with disease. Disease associated polymorphisms may include deletion or truncation of the gene, mutations that alter expression level, that affect the activity of the protein, *etc.*

Changes in the promoter or enhancer sequence that may affect expression levels of the gene can be compared to expression levels of the normal allele by various methods known in the art. Methods for determining promoter or enhancer strength include quantitation of the expressed natural protein; insertion of the variant control element into a vector with a

reporter gene such as β-galactosidase, luciferase, chloramphenicol acetyltransferase, *etc.* that provides for convenient quantitation; and the like.

A number of methods are available for analyzing nucleic acids for the presence of a specific sequence, *e.g.* a disease associated polymorphism. Where large amounts of DNA are available, genomic DNA is used directly. Alternatively, the region of interest is cloned into a suitable vector and grown in sufficient quantity for analysis. Cells that express the subject gene may be used as a source of mRNA, which may be assayed directly or reverse transcribed into cDNA for analysis. The nucleic acid may be amplified by conventional techniques, such as the polymerase chain reaction (PCR), to provide sufficient amounts for analysis. The use of the polymerase chain reaction is described in Saiki, *et al.* (1985), *Science* 239:487, and a review of techniques may be found in Sambrook, *et al.* Molecular Cloning: A Laboratory Manual, CSH Press 1989, pp.14.2B14.33. Alternatively, various methods are known in the art that utilize oligonucleotide ligation as a means of detecting polymorphisms, for examples see Riley *et al.* (1990), *Nucl. Acids Res.* 18:2887-2890; and Delahunty *et al.* (1996), *Am. J. Hum. Genet.* 58:1239-1246.

A detectable label may be included in an amplification reaction. Suitable labels include fluorochromes, *e.g.* fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE), 6-carboxy-X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), radioactive labels, *e.g.* ^{32}P, ^{35}S, ^{3}H; *etc.* The label may be a two stage system, where the amplified DNA is conjugated to biotin, haptens, *etc.* having a high affinity binding partner, *e.g.* avidin, specific antibodies, *etc.*, where the binding partner is conjugated to a detectable label. The label may be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label into the amplification product.

The sample nucleic acid, *e.g.* amplified or cloned fragment, is analyzed by one of a number of methods known in the art. The nucleic acid may be sequenced by dideoxy or other methods, and the sequence of bases compared to a wild-type sequence. Hybridization with the variant sequence may also be used to determine its presence, by Southern blots, dot blots, *etc.* The hybridization pattern of a control and variant sequence to an array of oligonucleotide probes immobilized on a solid support, as described in US 5,445,934, or in

WO 95/35505, may also be used as a means of detecting the presence of variant sequences. Single strand conformational polymorphism (SSCP) analysis, denaturing gradient gel electrophoresis (DGGE), and heteroduplex analysis in gel matrices are used to detect conformational changes created by DNA sequence variation as alterations in electrophoretic mobility. Alternatively, where a polymorphism creates or destroys a recognition site for a restriction endonuclease, the sample is digested with that endonuclease, and the products size fractionated to determine whether the fragment was digested. Fractionation is performed by gel or capillary electrophoresis, particularly acrylamide or agarose gels.

Screening for mutations may be based on the functional or antigenic characteristics of the protein. Protein truncation assays are useful in detecting deletions that may affect the biological activity of the protein. Various immunoassays designed to detect polymorphisms in the subject PDE interacting proteins may be used in screening. Where many diverse genetic mutations lead to a particular disease phenotype, functional protein assays have proven to be effective screening tools. The activity of the encoded protein may be determined by comparison with the wild-type protein.

Diagnostic methods of the subject invention in which the level of expression is of interest will typically involve comparison of the PDE interacting protein nucleic acid abundance of a sample of interest with that of a control value to determine any relative differences, where the difference may be measured qualitatively and/or quantitatively, which differences are then related to the presence or absence of an abnormal gene expression pattern. A variety of different methods for determining the nucleic acid abundance in a sample are known to those of skill in the art, where particular methods of interest include those described in: Pietu et al., Genome Res. (June 1996) 6: 492-503; Zhao et al., Gene (April 24, 1995) 156: 207-213; Soares , Curr. Opin. Biotechnol. (October 1997) 8: 542-546; Raval, J. Pharmacol Toxicol Methods (November 1994) 32: 125-127; Chalifour et al., Anal. Biochem (February 1, 1994) 216: 299-304; Stolz & Tuan, Mol. Biotechnol. (December 19960 6: 225-230; Hong et al., Bioscience Reports (1982) 2: 907; and McGraw, Anal. Biochem. (1984) 143: 298. Also of interest are the methods disclosed in WO 97/27317, the disclosure of which is herein incorporated by reference.

SCREENING ASSAYS

The subject PDE interacting proteins and polypeptides find use in various screening assays designed to identify therapeutic agents. The screening assays may be designed to identify agents that modulate, e.g. inhibit or enhance, the activity of the PDE interacting protein directly and thereby modulate the activity of the particular PDE that depends on the presence of the PDE interacting protein for its function. Alternatively, the assay may be designed to identify those agents that modify, e.g. enhance or inhibit, the activity of the PDE when present as a complex with the PDE interacting protein.

Of particular interest are screening methods that provide for qualitative/quantitative measurements of a PDE enzyme activity in the presence of a particular candidate therapeutic agent and its PDE interacting protein, as such screening methods are capable of identifying highly selective PDE modulatory, e.g. inhibitory, agents. For example, the assay could be an assay which measures the activity of a PDE interacting protein/enzyme complex in the presence and absence of a candidate inhibitor agent. In this preferred screening assay embodiment, the PDE interacting protein/PDE complex will generally be a naturally occurring complex, i.e. a complex between a cyclic nucleotide PDE and its naturally occurring PDE interacting protein partner. Of particular interest are complexes between a cAMP-PDEIV and a myomegalin protein.

The screening method may be an *in vitro* or *in vivo* format, where both formats are readily developed by those of skill in the art. Depending on the particular method, one or more of, usually one of, the components of the screening assay may be labeled, where by labeled is meant that the components comprise a detectable moiety, *e.g.* a fluorescent or radioactive tag, or a member of a signal producing system, *e.g.* biotin for binding to an enzyme-streptavidin conjugate in which the enzyme is capable of converting a substrate to a chromogenic product.

A variety of other reagents may be included in the screening assay. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc. that are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, *etc.* may be used. Specific PDE activity assays of interest include those described in U.S. Patent Nos. 5,798,373 and 5,580,888, the disclosures of which are herein incorporated by reference.

A variety of different candidate agents may be screened by the above methods. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, *etc.* to produce structural analogs.

PDE INTERACTING PROTEIN NUCLEIC ACID AND POLYPEPTIDE THERAPEUTIC COMPOSITIONS

The nucleic acid compositions of the subject invention also find use as therapeutic agents in situations where one wishes to enhance the PDE interacting protein activity in a host, e.g. in a mammalian host in which PDE interacting protein activity is sufficiently low such that a disease condition is present, etc. The PDE interacting protein genes, gene fragments, or the encoded proteins or protein fragments are useful in gene therapy to treat disorders associated with defects the PDE interacting protein gene expression. Expression vectors may be used to introduce the gene into a cell. Such vectors generally have convenient restriction sites located near the promoter sequence to provide for the insertion of nucleic acid sequences. Transcription cassettes may be prepared comprising a transcription initiation

region, the target gene or fragment thereof, and a transcriptional termination region. The transcription cassettes may be introduced into a variety of vectors, *e.g.* plasmid; retrovirus, *e.g.* lentivirus; adenovirus; and the like, where the vectors are able to transiently or stably be maintained in the cells, usually for a period of at least about one day, more usually for a period of at least about several days to several weeks.

The gene or protein may be introduced into tissues or host cells by any number of routes, including viral infection, microinjection, or fusion of vesicles. Jet injection may also be used for intramuscular administration, as described by Furth *et al.* (1992), *Anal Biochem* **205**:365-368. The DNA may be coated onto gold microparticles, and delivered intradermally by a particle bombardment device, or "gene gun" as described in the literature (see, for example, Tang *et al.* (1992), *Nature* **356**:152-154), where gold microprojectiles are coated with the DNA, then bombarded into skin cells.

METHODS OF MODULATING PDE INTERACTING PROTEIN ACTIVITY IN A HOST

Also provided are methods of regulating, including enhancing and inhibiting, PDE interacting protein activity in a host. Where the PDE interacting protein activity occurs *in vivo* in a host, an effective amount of active agent that modulates the activity, e.g. reduces the activity, of the PDE interacting protein *in vivo* (e.g. the activity of the naturally occurring PDE/interacting protein complex), is administered to the host. The active agent may be a variety of different compounds, including a naturally occurring or synthetic small molecule compound, an antibody, fragment or derivative thereof, an antisense composition, and the like.

Naturally occurring or synthetic small molecule compounds of interest include numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof.

Also of interest as active agent are antibodies that modulate, e.g. reduce, if not inhibit, PDE interacting protein activity in the host. Suitable antibodies are obtained by immunizing a host animal with peptides comprising all or a portion of the subject proteins, such as found in the polypeptide compositions of the subject invention. Suitable host animals include mouse, rat sheep, goat, hamster, rabbit, *etc.* The origin of the protein immunogen may be mouse, human, rat, monkey *etc.* The host animal will generally be a different species than the immunogen, *e.g.* human protein used to immunize mice, *etc.*

The immunogen may comprise the complete protein, or fragments and derivatives thereof. Preferred immunogens comprise all or a part of the PDE interacting protein, where these residues contain the post-translation modifications, such as glycosylation, found on the native protein. Immunogens comprising the extracellular domain are produced in a variety of ways known in the art, *e.g.* expression of cloned genes using conventional recombinant methods, isolation from HEC, *etc.*

For preparation of polyclonal antibodies, the first step is immunization of the host animal with the immunogen, where the immunogen will preferably be in substantially pure form, comprising less than about 1% contaminant. The immunogen may comprise complete PDE interacting protein, fragments or derivatives thereof. To increase the immune response of the host animal, the protein or peptide may be combined with an adjuvant, where suitable adjuvants include alum, dextran, sulfate, large polymeric anions, oil & water emulsions, e.g. Freund's adjuvant, Freund's complete adjuvant, and the like. The immunogen may also be conjugated to synthetic carrier proteins or synthetic antigens. A variety of hosts may be immunized to produce the polyclonal antibodies. Such hosts include rabbits, guinea pigs, rodents, e.g. mice, rats, sheep, goats, and the like. The immunogen is administered to the host, usually intradermally, with an initial dosage followed by one or more, usually at least two, additional booster dosages. Following immunization, the blood from the host will be collected, followed by separation of the serum from the blood cells. The Ig present in the resultant antiserum may be further fractionated using known methods, such as ammonium salt fractionation, DEAE chromatography, and the like.

Monoclonal antibodies are produced by conventional techniques. Generally, the spleen and/or lymph nodes of an immunized host animal provide a source of plasma cells. The plasma cells are immortalized by fusion with myeloma cells to produce hybridoma cells. Culture supernatant from individual hybridomas is screened using standard techniques to

identify those producing antibodies with the desired specificity. Suitable animals for production of monoclonal antibodies to the human protein include mouse, rat, hamster, *etc.* To raise antibodies against the mouse protein, the animal will generally be a hamster, guinea pig, rabbit, etc. The antibody may be purified from the hybridoma cell supernatants or ascites fluid by conventional techniques, *e.g.* affinity chromatography using PDE-interacting protein bound to an insoluble support, protein A sepharose, *etc.*

The antibody may be produced as a single chain, instead of the normal multimeric structure. Single chain antibodies are described in Jost *et al.* (1994) J.B.C. 269:26267-73, and others. DNA sequences encoding the variable region of the heavy chain and the variable region of the light chain are ligated to a spacer encoding at least about 4 amino acids of small neutral amino acids, including glycine and/or serine. The protein encoded by this fusion allows assembly of a functional variable region that retains the specificity and affinity of the original antibody.

For *in vivo* use, particularly for injection into humans, it is desirable to decrease the antigenicity of the antibody. An immune response of a recipient against the blocking agent will potentially decrease the period of time that the therapy is effective. Methods of humanizing antibodies are known in the art. The humanized antibody may be the product of an animal having transgenic human immunoglobulin constant region genes (see for example International Patent Applications WO 90/10077 and WO 90/04036). Alternatively, the antibody of interest may be engineered by recombinant DNA techniques to substitute the CH1, CH2, CH3, hinge domains, and/or the framework domain with the corresponding human sequence (see WO 92/02190).

The use of Ig cDNA for construction of chimeric immunoglobulin genes is known in the art (Liu *et al.* (1987) P.N.A.S. 84:3439 and (1987) J. Immunol. 139:3521). mRNA is isolated from a hybridoma or other cell producing the antibody and used to produce cDNA. The cDNA of interest may be amplified by the polymerase chain reaction using specific primers (U.S. Patent nos. 4,683,195 and 4,683,202). Alternatively, a library is made and screened to isolate the sequence of interest. The DNA sequence encoding the variable region of the antibody is then fused to human constant region sequences. The sequences of human constant regions genes may be found in Kabat *et al.* (1991) Sequences of Proteins of Immunological Interest, N.I.H. publication no. 91-3242. Human C region genes are readily available from known clones. The choice of isotype will be guided by the desired effector

functions, such as complement fixation, or activity in antibody-dependent cellular cytotoxicity. Preferred isotypes are IgG1, IgG3 and IgG4. Either of the human light chain constant regions, kappa or lambda, may be used. The chimeric, humanized antibody is then expressed by conventional methods.

Antibody fragments, such as Fv, $F(ab)_2$ and Fab may be prepared by cleavage of the intact protein, *e.g.* by protease or chemical cleavage. Alternatively, a truncated gene is designed. For example, a chimeric gene encoding a portion of the $F(ab)_2$ fragment would include DNA sequences encoding the CH1 domain and hinge region of the H chain, followed by a translational stop codon to yield the truncated molecule.

Consensus sequences of H and L J regions may be used to design oligonucleotides for use as primers to introduce useful restriction sites into the J region for subsequent linkage of V region segments to human C region segments. C region cDNA can be modified by site directed mutagenesis to place a restriction site at the analogous position in the human sequence.

Expression vectors include plasmids, retroviruses, YACs, EBV derived episomes, and the like. A convenient vector is one that encodes a functionally complete human CH or CL immunoglobulin sequence, with appropriate restriction sites engineered so that any VH or VL sequence can be easily inserted and expressed. In such vectors, splicing usually occurs between the splice donor site in the inserted J region and the splice acceptor site preceding the human C region, and also at the splice regions that occur within the human CH exons. Polyadenylation and transcription termination occur at native chromosomal sites downstream of the coding regions. The resulting chimeric antibody may be joined to any strong promoter, including retroviral LTRs, *e.g.* SV-40 early promoter, (Okayama *et al.* (1983) Mol. Cell. Bio. 3:280), Rous sarcoma virus LTR (Gorman *et al.* (1982) P.N.A.S. 79:6777), and moloney murine leukemia virus LTR (Grosschedl *et al.* (1985) Cell 41:885); native Ig promoters, *etc.*

In yet other embodiments of the invention, the active agent is an agent that modulates, and generally decreases or down regulates, the expression of the gene in the host. Antisense molecules can be used to down-regulate expression of the protein in cells. The anti-sense reagent may be antisense oligonucleotides (ODN), particularly synthetic ODN having chemical modifications from native nucleic acids, or nucleic acid constructs that express such anti-sense molecules as RNA. The antisense sequence is complementary to the mRNA of the targeted gene, and inhibits expression of the targeted gene products. Antisense molecules

inhibit gene expression through various mechanisms, *e.g.* by reducing the amount of mRNA available for translation, through activation of RNAse H, or steric hindrance. One or a combination of antisense molecules may be administered, where a combination may comprise multiple different sequences.

Antisense molecules may be produced by expression of all or a part of the target gene sequence in an appropriate vector, where the transcriptional initiation is oriented such that an antisense strand is produced as an RNA molecule. Alternatively, the antisense molecule is a synthetic oligonucleotide. Antisense oligonucleotides will generally be at least about 7, usually at least about 12, more usually at least about 20 nucleotides in length, and not more than about 500, usually not more than about 50, more usually not more than about 35 nucleotides in length, where the length is governed by efficiency of inhibition, specificity, including absence of cross-reactivity, and the like. It has been found that short oligonucleotides, of from 7 to 8 bases in length, can be strong and selective inhibitors of gene expression (see Wagner *et al.* (1996), *Nature Biotechnol.* 14:840-844).

A specific region or regions of the endogenous sense strand mRNA sequence is chosen to be complemented by the antisense sequence. Selection of a specific sequence for the oligonucleotide may use an empirical method, where several candidate sequences are assayed for inhibition of expression of the target gene in an *in vitro* or animal model. A combination of sequences may also be used, where several regions of the mRNA sequence are selected for antisense complementation.

Antisense oligonucleotides may be chemically synthesized by methods known in the art (see Wagner *et al.* (1993), *supra*, and Milligan *et al.*, *supra*.) Preferred oligonucleotides are chemically modified from the native phosphodiester structure, in order to increase their intracellular stability and binding affinity. A number of such modifications have been described in the literature, which alter the chemistry of the backbone, sugars or heterocyclic bases.

Among useful changes in the backbone chemistry are phosphorothioates; phosphorodithioates, where both of the non-bridging oxygens are substituted with sulfur; phosphoroamidites; alkyl phosphotriesters and boranophosphates. Achiral phosphate derivatives include 3′-O′-5′-S-phosphorothioate, 3′-S-5′-O-phosphorothioate, 3′-CH_2-5′-O-phosphonate and 3′-NH-5′-O-phosphoroamidate. Peptide nucleic acids replace the entire ribose phosphodiester backbone with a peptide linkage. Sugar modifications are also used to

enhance stability and affinity. The α-anomer of deoxyribose may be used, where the base is inverted with respect to the natural β-anomer. The 2′-OH of the ribose sugar may be altered to form 2′-O-methyl or 2′-O-allyl sugars, which provides resistance to degradation without comprising affinity. Modification of the heterocyclic bases must maintain proper base pairing. Some useful substitutions include deoxyuridine for deoxythymidine; 5-methyl-2′-deoxycytidine and 5-bromo-2′-deoxycytidine for deoxycytidine. 5- propynyl-2′-deoxyuridine and 5-propynyl-2′-deoxycytidine have been shown to increase affinity and biological activity when substituted for deoxythymidine and deoxycytidine, respectively.

As an alternative to anti-sense inhibitors, catalytic nucleic acid compounds, *e.g.* ribozymes, anti-sense conjugates, *etc.* may be used to inhibit gene expression. Ribozymes may be synthesized *in vitro* and administered to the patient, or may be encoded on an expression vector, from which the ribozyme is synthesized in the targeted cell (for example, see International patent application WO 9523225, and Beigelman *et al.* (1995), *Nucl. Acids Res.* **23**:4434-42). Examples of oligonucleotides with catalytic activity are described in WO 9506764. Conjugates of anti-sense ODN with a metal complex, *e.g.* terpyridylCu(II), capable of mediating mRNA hydrolysis are described in Bashkin *et al.* (1995), *Appl. Biochem. Biotechnol.* **54**:43-56.

As mentioned above, an effective amount of the active agent is administered to the host, where "effective amount" means a dosage sufficient to produce a desired result, where the desired result in the desired modulation, e.g. enhancement, reduction, of PDE interacting protein activity, which in turn leads to a desired effect on the state of the disease condition being treated, e.g. a reduction in the level of inflammation, etc.

In the subject methods, the active agent(s) may be administered to the host using any convenient means capable of resulting in the desired inhibition of PDE interacting protein activity. Thus, the agent can be incorporated into a variety of formulations for therapeutic administration. More particularly, the agents of the present invention can be formulated into pharmaceutical compositions by combination with appropriate, pharmaceutically acceptable carriers or diluents, and may be formulated into preparations in solid, semi-solid, liquid or gaseous forms, such as tablets, capsules, powders, granules, ointments, solutions, suppositories, injections, inhalants and aerosols.

As such, administration of the agents can be achieved in various ways, including oral, buccal, rectal, parenteral, intraperitoneal, intradermal, transdermal, intracheal,etc., administration.

In pharmaceutical dosage forms, the agents may be administered in the form of their pharmaceutically acceptable salts, or they may also be used alone or in appropriate association, as well as in combination, with other pharmaceutically active compounds. The following methods and excipients are merely exemplary and are in no way limiting.

For oral preparations, the agents can be used alone or in combination with appropriate additives to make tablets, powders, granules or capsules, for example, with conventional additives, such as lactose, mannitol, corn starch or potato starch; with binders, such as crystalline cellulose, cellulose derivatives, acacia, corn starch or gelatins; with disintegrators, such as corn starch, potato starch or sodium carboxymethylcellulose; with lubricants, such as talc or magnesium stearate; and if desired, with diluents, buffering agents, moistening agents, preservatives and flavoring agents.

The agents can be formulated into preparations for injection by dissolving, suspending or emulsifying them in an aqueous or nonaqueous solvent, such as vegetable or other similar oils, synthetic aliphatic acid glycerides, esters of higher aliphatic acids or propylene glycol; and if desired, with conventional additives such as solubilizers, isotonic agents, suspending agents, emulsifying agents, stabilizers and preservatives.

The agents can be utilized in aerosol formulation to be administered via inhalation. The compounds of the present invention can be formulated into pressurized acceptable propellants such as dichlorodifluoromethane, propane, nitrogen and the like.

Furthermore, the agents can be made into suppositories by mixing with a variety of bases such as emulsifying bases or water-soluble bases. The compounds of the present invention can be administered rectally via a suppository. The suppository can include vehicles such as cocoa butter, carbowaxes and polyethylene glycols, which melt at body temperature, yet are solidified at room temperature.

Unit dosage forms for oral or rectal administration such as syrups, elixirs, and suspensions may be provided wherein each dosage unit, for example, teaspoonful, tablespoonful, tablet or suppository, contains a predetermined amount of the composition containing one or more inhibitors. Similarly, unit dosage forms for injection or intravenous

administration may comprise the inhibitor(s) in a composition as a solution in sterile water, normal saline or another pharmaceutically acceptable carrier.

The term "unit dosage form," as used herein, refers to physically discrete units suitable as unitary dosages for human and animal subjects, each unit containing a predetermined quantity of compounds of the present invention calculated in an amount sufficient to produce the desired effect in association with a pharmaceutically acceptable diluent, carrier or vehicle. The specifications for the novel unit dosage forms of the present invention depend on the particular compound employed and the effect to be achieved, and the pharmacodynamics associated with each compound in the host.

The pharmaceutically acceptable excipients, such as vehicles, adjuvants, carriers or diluents, are readily available to the public. Moreover, pharmaceutically acceptable auxiliary substances, such as pH adjusting and buffering agents, tonicity adjusting agents, stabilizers, wetting agents and the like, are readily available to the public.

Where the agent is a polypeptide, polynucleotide, analog or mimetic thereof, e.g. antisense composition, it may be introduced into tissues or host cells by any number of routes, including viral infection, microinjection, or fusion of vesicles. Jet injection may also be used for intramuscular administration, as described by Furth *et al.* (1992), *Anal Biochem* **205**:365-368. The DNA may be coated onto gold microparticles, and delivered intradermally by a particle bombardment device, or "gene gun" as described in the literature (see, for example, Tang *et al.* (1992), *Nature* **356**:152-154), where gold microprojectiles are coated with the DNA, then bombarded into skin cells.

Those of skill in the art will readily appreciate that dose levels can vary as a function of the specific compound, the severity of the symptoms and the susceptibility of the subject to side effects. Preferred dosages for a given compound are readily determinable by those of skill in the art by a variety of means.

The subject methods find use in the treatment of a variety of different disease conditions involving PDE interacting protein activity, particularly in those disease conditions in which the selective inhibition of PDE activity, more particularly PDEIV activity, results in treatment of the disease condition where targeting of the PDE interacting protein by the therapeutic agent results in modulated, e.g. reduced or enhanced, activity of its corresponding PDE.

Specific disease of interest as treatable by the subject methods include: asthma, including inflamed lung associate asthma, cystic fibrosis, inflammatory airway disease, chronic bronchitis, eosinophilic granuloma, psoriasis and other benign and malignant proliferative skin diseases, endotoxic shock, septic shock, ulcerative colitis, Crohn's disease, reperfusion injury, or the myocardium and brain, inflammatory arthritis, chronic gloerulonephritis, atopic dermatitis, urticaria, adult respiratory distress syndrome, diabetes insipidus, allergic rhinitis, allergic conjunctivitis, vernal conjunctivitis, arterial restinosis and artherosclerosis, inflammatory diseases associated with irritation and pain, rheumatoid arthritis, ankylosing spondylitis, transplant rejection and graft versus host disease, disease conditions associated with hypersecretion of gastric acid, disease conditions in which cytokines are mediators, e.g. sepsis, and septic shock, and the like.

By treatment is meant at least an amelioration of the symptoms associated with the pathological condition afflicting the host, where amelioration is used in a broad sense to refer to at least a reduction in the magnitude of a parameter, *e.g.* symptom, associated with the pathological condition being treated, such as inflammation, etc. As such, treatment also includes situations where the pathological condition, or at least symptoms associated therewith, are completely inhibited, *e.g.* prevented from happening, or stopped, *e.g.* terminated, such that the host no longer suffers from the pathological condition, or at least the symptoms that characterize the pathological condition.

A variety of hosts are treatable according to the subject methods. Generally such hosts are "mammals" or "mammalian," where these terms are used broadly to describe organisms which are within the class mammalia, including the orders carnivore (*e.g.*, dogs and cats), rodentia (*e.g.*, mice, guinea pigs, and rats), and primates (*e.g.*, humans, chimpanzees, and monkeys). In many embodiments, the hosts will be humans.

Kits with unit doses of the active agent, usually in oral or injectable doses, are provided. In such kits, in addition to the containers containing the unit doses will be an informational package insert describing the use and attendant benefits of the drugs in treating pathological condition of interest. Preferred compounds and unit doses are those described herein above.

The following examples are offered primarily for purposes of illustration. It will be readily apparent to those skilled in the art that the formulations, dosages, methods of

administration, and other parameters of this invention may be further modified or substituted in various ways without departing from the spirit and scope of the invention.

EXPERIMENTAL

I. Screening of the yeast two hybrid system cDNA brain library

To identify proteins that interact with a PDE4, cDNA coding for the amino terminus of PDE4D3 or for a region corresponding to a.a. 114-672 were inserted into pGBT9 vectors and used for screening of a Matchmaker rat brain library subcloned in pGAD10 vector (Clontech, Palo Alto, CA). The fragment encoding the autoinhibitory (UCR2), catalytic, and carboxy terminal domains of rPDE4D3 (aa 114-672) was amplified by PCR with the full-length cDNA using the following forward and reverse primers with incorporated restriction sites and Stop codon. EcoRI: 5' CGG AAT TCG AGG AGG CCT ACC AGA AAC 3' (GUPA4) (SEQ ID NO:06) and SalI/TAG: 5' TGA GTC GAC TAC GTG TCA AGG CAA CAA TGG TC 3' (GUPA3) (SEQ ID NO:07). The PCR products were cloned into EcoRI/SalI site of pGBT9 (Clontech) downstream of the Gal4 activation domain. The PCR was performed in presence of recombinant Pfu polymerase (Stratagene) at low cycle number (10 cycles) to ensure high fidelity reading. The insertions were entirely sequenced to confirm the correct reading frame and the sequence. Sequencing was performed by the Molecular Biology facility at Stanford University using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq DNA Polymerase, FS (Perkin Elmer).

Of the positive clones isolated from the screening of the rat brain library, 187 gave strong positive signal while 81 gave only a weak signal. Of the strong positive clones, PBP46 was further characterized. This clone contained an insert of aprroximately 2.8 kb. The interaction of the clone with the PDE was confirmed by subcloning the cDNA fragment in both pGBT9 and pGAD10 and by testing growth and β-galactosidase activity in the yeast two hybrid system. The clone continued to show strong interaction with the 1.6 fragment of PDE4D3.

II. Screening for the full length myomegalin clone

A homology search (BLAST) using the sequence of PBP46 clone showed no significant identity to sequences in any public domain database. This clone was then used to probe a blot with RNA from multiple tissues. A transcript of approximately 8.0-8.5 kb

hybridized to the probe in several tissues, the highest level of expression being observed in the rat skeletal muscle and heart. Lower levels of expression were detected in brain, liver and lung. In the testis a transcript of 2.0-2.4 kb was consistently observed. The expression in the testis was confirmed by PCR and by screening a rat testis library. Two clones containing the 3' end sequence of myomegalin were retrieved from this library.

To obtain the complete sequence of the 8.0-8.5 transcript, a rat skeletal muscle cDNA library was screened with the PBP46 cDNA. From this screening, 2 clones were retrieved. However, the clones did not yield a complete ORF. Screening was then repeated six more times with oligonucleotides corresponding to the 5' end of the longest clones. From this multiple screening, 21 overlapping clones were obtained. Merging of the sequences from the different clones yielded a 9 kb sequence, a size in agreement with the size of the transcript derived from rat heart and skeletal muscles. See Fig. 2. Conceptual translation of the nucleotide sequence uncovered an open reading frame of a protein of 2324 amino acids corresponding to a calculated MW of 261 kDa. See Fig. 1.

To analyze tissue distribution of the rat myomegalin transcripts, Northern blot analysis was performed using radioactively labeled probes corresponding to the 3' end (probe 1; 1000 bp) and the 5' end (probe 2; 665 bp) of the myomegalin open reading frame. Transcripts of various sizes were found in various tissues using either probe 1 or probe 2 or both. The results indicated that there are at least four different transcripts of rat myomegalin: two expressed in heart (7.5 and 5.9 kb); two in skeletal muscle (7.5 and 4.3 kb) and one in testis (2.5 kb). The 2.5 kb variant roughly corresponds to the PBP46 clone, and is expressed exclusively in rat testis.

III. Screening of the EST/database

To determine whether mouse or human sequences analogous to the rat myomegalin are present in public domain databases, the rat sequence was used for a BLAST search of GenBank and EST libraries. The following EST were retrieved. AA755885, AA110441, W23471, AA333456, AA489265. These sequences are more than 90% homologous to the rat sequence. Sequence AL021920 contains a genomic fragment from human chromosome 1p35.1-p36.21. Several exons overlap with the rat sequence from residue 1215 until residue 1444. Thus myomegalin must reside on human chromosome 1p35.1-p36. KIAA0454 (accession # AB007923), KIAA0477 (acccesion # AB007946) are two clones containing

portion of the human myomegalin sequence since they are more than 90% homologous to the rat ORF. These human clones were merged to obtain a full length human sequence homologous to myomegalin. See Fig. 4. The human open reading frame coded for a protein of 2517 residues and a calculated molecular weight of 282.1 kDa. See Fig. 5.

Alignment of the human and rat sequence showed identity from aa 235 of rat myomegalin to the end. In the amino terminus region, the two sequences showed only weak homologies. The reason for this discrepancy is at present unclear. It is possible that it is due to species differences. The junction where the rat sequence diverges from the human was derived from four clones isolated from the rat skeletal muscle library, lessening the possibility that cloning artifact is at the basis of this discrepancy. The presence of the junction was further confirmed by PCR analysis of rat heart mRNA (data not shown). However, further blast searches with the region encompassing the 5' end of myomegalin did not yield mouse EST fragments overlapping the junction. Conversely, several EST clones confirming the human junction were retrieved from human and mouse EST databases.

IV. Protein/protein interaction

Several attempts were made to confirm the interaction between myomegalin and PDE4D3. However, due to the insolubility of the full length or truncated myomegalin immunoprecipitation experiments could not be performed. In an alternative approach, PBP46 was cotransfected with PDE4D3 in COS 7 cells and the PDE activity was determined in the particulate fraction of the cell. If PDE4D3 interacts with PBP46, an increase in the particulate PDE activity would be expected. Two to three fold increase in the particulate PDE4D3 activity was detected when plasmids containing PBP46 and PDE4D3 were cotransfected in COS7 cells.

V. Subcellular localization of myomegalin

To investigate the subcellular localization of myomegalin the PBP46 clone was subcloned in frame to a flag tag and expressed in COS7 cells. The recombinant protein thus obtained was entirely recovered in the particulatefraction and could be extracted only with buffer containing SDS. Expression in transfected cells was further assessed by immunofluorescence (IF) using the flag antibody. The flag tagged recombinant protein

encoded in PBP46 was entirely localized in the Golgi/centrosomal region of COS7 cells. No attempts were made to express the full-length myomegalin cDNA.

VI. Western blot analysis of muscle and testis extracts

Polyclonal antibodies were raised in rabbit against peptides corresponding to the carboxyl terminus region of myomegalin. These antibodies recognize in testis a protein of approximately 64 kDa. In heart and muscle, proteins of 280,250 and 200 kDa were observed. It is at present unknown whether these are native proteins or products of proteolysis. When these antibodies were used for IF localization. a region corresponding to the Golgi/centrosomal region is intensely labeled.

It is apparent from the above results and discussion that polynucleotides encoding novel mammalian PDE interacting proteins, such as myomegalin, as well as the novel polypeptides encoded thereby, are provided. The subject invention is important for both research and therapeutic applications. For example, identification of the subject PDE interacting proteins provides for the ability to screen potential PDE inhibitors with PDE/PDE interacting protein complexes, where the results of such screening procedures should be more indicative of *in vivo* activity of a potential agent than screening procedures in which PDE is used by itself. Accordingly, the subject invention provides for a significant contribution to the art.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

WHAT IS CLAIMED IS:

1. A polynucleotide present in other than its natural environment encoding a PDE interacting polypeptide.

2. The polynucleotide according to Claim 1, wherein said polynucleotide encodes a myomegalin protein.

3. A fragment of a polynucleotide according to Claim 1.

4. An PDE interacting polypeptide present in other than its naturally occurring environment.

5. The polypeptide according to Claim 4, wherein said polypeptide is a myomegalin protein.

6. A fragment of a polypeptide according to Claim 4.

7. Substantially pure PDE interacting protein.

8. Isolated PDE interacting protein.

9. An expression cassette comprising a transcriptional initiation region functional in an expression host, a polynucleotide having a nucleotide sequence found in the nucleic acid according to Claim 1 under the transcriptional regulation of said transcriptional initiation region, and a transcriptional termination region functional in said expression host.

10. A cell comprising an expression cassette according to Claim 9 as part of an extrachromosomal element or integrated into the genome of a host cell as a result of introduction of said expression cassette into said host cell.

11. The cellular progeny of the cell according to Claim 10.

12. A method of producing an PDE interacting polypeptide, said method comprising:

growing a cell according to Claim 10, whereby said polypeptide is expressed; and

isolating said polypeptide substantially free of other proteins.

13. A monoclonal antibody binding specifically to a PDE interacting protein.

14. The monoclonal antibody according to Claim 13, wherein said antibody inhibits the activity of at least one of PDE or a PDE interacting protein.

15. The monoclonal antibody according to Claim 13, wherein said antibody is a humanized antibody.

16. A method of determining whether an agent modulates the activity of a PDE, said method comprising:

contacting a complex of said PDE and a PDE interacting protein with said agent; and

determining the effect of said agent on the activity of said PDE.

17. The method according to Claim 16, wherein said agent is a small molecule.

18. The method according to Claim 16, wherein said agent is an antibody.

19. The method according to Claim 18, wherein said agent is a monoclonal antibody.

20. A method for modulating the activity of a PDE interacting protein, said method comprising:

contacting said PDE interacting protein with an agent that modulates the activity of said PDE interacting protein.

FIG. 1

>myomegalin protein
MSNGYRTLSQHLNDLKKENFSLKLRIYFLEERMQQKYEVSREDVYKRNIELKVEVESLKRELQDRKQHL
HKTWADEEDLNSQNEAELRRQVEEPQQETEHVYELLDNNIQLLQEESRFAKDEATQMETLVEAEKGCNL
ELSERWKDATKNREDAPGDQVKLDQYSAALAQRDRRIEELRQSLAAQEGLVEQLSREKQQLLHLLEEPG
GMEVQPMPKGLPTQQKPDLNETPTTQPSVSDSHLAELQDKIQQTEVTNKILQEKLNDMSCELRSAQESS
QKQDTTIQSLKEMLKSRESETEELYQVIEGQNDTMAKLPEMLHQSQLGQLQSSEGIAPAQQQVALLDLQ
SALFCSQLEIQKLQRLLRQKERQLADGKRCMQFVEAAAQEREQQKEAAWKHNQELRKALQHLQGELHSK
SQQLHVLEAEKYNEIRTQGQNIQHLSHSLSHKEQLIQELQELLQYRDTTDKTLDTNEVFLEKLRQRIQD
RAVALERVIDEKFSALEEKDKELRQLRLAVRDRDHDLERLRCVLSANEATMQSMESLLRARGLEVEQLI
ATCQNLQWLKEELETKFGHWQKEQESIIQQLQTSLHDRNKEVEDLSATLLHKLGPGQSEVAEELCQRLQ
RKERVLQDLLSDRNKQAMEHEMEVQGLLQSMGTREQERQAVAEKMVQAFMERNSELQALRQYLGGKELM
AASQAFISNQPAGATSVGPHHGEQTDQGSTQMPSRDDSTSLTAREEASIPRSTLGDSDTVAGLEKELSN
AKEELELMAKKERESQIELSALQSMMAVQEEELQVQAADLESLTRNIQIKEDLIKDLQMQLVDPEDMPA
MERLTQEVLLLREKVASVEPQGQEGSENRRQQLLLMLEGLVDERSRLNEALQAERQLYSSLVKFHAQPE
ISERDRTLQVELEGAQVLRSRLEEVLGRSLERLSRLETLAAIGGATAGDETEDTSTQFTDSIEEEAAHN
SHQQLIKVSLEKSLTTMETQNTCLQPPSPVGEDGNRHLQEEMLHLRAEIHQPLEEKRKAEAELKELKAQ
IEEAGFSSVSHIRNTMLSLCLCLENAELKEQMGEAMSDGWEVEEDKEKGEVMVETVVAKGGLSEDSLQA
EFRKVQGRLKSAYNIINLLKEQLVLRSSEGNTKEMPEFLVRLAREVDRMNMGLPSSEKHQHQEQENMTA
RPGPRPQSLKLGTALSVDGYQLENKSQAQDSGHQPEFSLPGSTKHLRSQLAQCRQRYQDLQEKLLISEA
TVFAQANQLEKYRAILSESLVKQDSKQIQVDLQDLGYETCGRSENEAEREETTSPECEEHGNLKPVVLV
EGLCSEQGYLDPVLVSSPVKNPWRTSQEARRIQAQGTSDNSSLLRKDIRNLKAQLPNAYKVLQNLRSRV
RSLSATSDYSSSLERPRKLIAVATLEGASPHSVTDEDEGLLSDGTGAFYPPGLQAKKNLENLIQRVSQL
EAQLPKTGLEGKLAEELKSASWPGKYDSLIQDQARKTVISASENTKREKDLFSSHPTFERYVKSFEDLL
RNNDLTTYLGQSFREQLSSRRSVTDRLTSKFSTKDHKSEKEEVGLEPLAFRFSRELQEKEKVIEVLQAK
VDTRFFSPPSSHAASESHRCASSTSFLSDDIEACSDMDVASEYTHYEEKKPSPSNSAASASQGLKGEPR
SSSISLPTPQNPPKEASQAQPGFHFNSIPKPASLSQAPMHFTVPSFMPFGPSGPPLLGCCETPVVSLAE
AQQELQMLQKQLGRSVSIAPPTSTSTLLSNHTEASSPRYSNPAQPHSPARGTIELGRILEPGYLGSGQW
DMMRPQKGSISGELSSGSSMYQLNSKPTGADLLEEHLGEIRNLRQRLEESICVNDRLREQLQHRLSSTA
RENGSTSHFYSQGLESMPQLYNENRALREENQSLQTRLSHASRGHSQEVDHLREALLSSSSQLQELEKE
LEQQKAERRQLLEDLQEKQDEIVHFREERLSLQENNSRLQHKLALLQQQCEEKQQLSLSLQSELQIYES
LYENPKKGLKAFSLDSCYQVPGELSCLVAEIRALRVQLEQSIQVNNRLRLQLEQQMDHGAGKASLSSCP
VNQSFSAKAELANQQPPFQGSAASPPVRDVGLNSPPVVLPSNSCSVPGSDSAIISRTNNGSDESAATKT
PPKMEVDAADGPFASGHGRHVIGHVDDYDALQQQIGEGKLLIQKILSLTRPARSVPALDAQGTEAPGTK
SVHELRSSARALNHSLEESASLLTMFWRAALPNSHGSVLVGEEGNLMEKELLDLRAQVSQQQQLLQSTA
VRLKTANQRKKSMEQFIVSHLTRTHDVLKKARTNLEMKSFRALMCTPAL (SEQ ID NO:01)

FIG. 2 >MYOMEGALIN complete DNA

CCGGTCCCCTTTGGTAGTAGTATCTCAGAGCTCGCCCCATAGTTTCATAGTTCATGTCTGGTTTGTTCT
TATGCTTTCCCCAGAGCTTCGAGACAGCCTTTGAGTCCACCAGCTTGAATATGCCCTTTTCTCTCTGAG
TCCATTTAATATACCTGGGACAAGTATTTTTATCTTGAAGCAGATCTAAAAGAAACTCCCACAGATAGG
TTGTGTTTCCTTTTCCTTCTCTGGCTTTCTTCTTGACTCCTAACTCAGGAGACCCATTGGAAACTGGTG
ACTGCTGGGTCTTTGGTTTACGGCCAACTTTCTTCTTTTTCATTGGTTCGTGGCTGTCTGGTGAAGTAT
GGATAGGCGAGGCATCCATTGGTTCAGACTCCTCTGTTGACACCTCCACTACAGTCTCCGTAATGACAT
CTGGCCTCATCGCAGCATGGATAAAATCGGATTCTTGAATCCTCAAGCAGGTAGGAGACTCCATATGAA
GCAGGGCTTCAGCAGCTTCAATGGTCTTATCCGTACAGTGTGCATTACTGCTGTGAACTGATGCTTCCA
CGGCCCGGATGAGCAGATCCAGCTGGTTCGTGGGTCCCTCATGCAGAGACGTCGCCATGTTTATCCCGG
GGCTGGAACTGCTGCTTCACATTGACTTACACCCTGAGCAGCGGCGACAGGGGAGAAGGCGGAACCCGC
GGCCGGAGACACACGCCGTGCGGGCGGCACACACTCACGCACTCGCACACACTCCGACGCCCGGATCCT
TGCGCGTCCTCCGACAGGAAGCGGCGGCCGGCCCGCCTCCCGCCGCCGGGCTGAGCAGCCCCACCACCT
AACGGCAGGGGCGGCGGCGGCCCCGGCTGGCAACGCGATCCTTCCGCCCCGCGCCCAGACAGGAAGTCC
CGGGCGCCGGCAGCCAGCGGCCGCACGGACACCTGAGGCTGGGGAGCCCGCAGGCCGCCCTCGGGGACG
CGGGCCTCGGCAGGAAAAGGCGCGCTTCACGTTCTGCGGAAGCGAAGTCTGCAAATGTCCCCTCAGCAT
GGTCTTCCTCCTGGCTCAATCTGTCTCACCTTCAGGTGATCCTAGGACTGGGGCTCCTTTCCAGGTCCC
CAGTTTCTCAAGTCGATCTTCTACCTCCCTCTTGATTTTCTACTCCATTGCTGGAAAGCTCCAGAACAG
AGCCTCCGCCGCCAACCACTGCTGATGCCATCGCGTCTTCCCTGAGCAAGTTTCGAACGCTGCGAATCA
ATGTAATTACGGCTCAGATGATTGCCAGGGTTATCGGTTTCATGTTCTAATTCAATAGTGATGGAGTAG
ACATCCAGAAGTCCAGTCTTCTAAAGATGATTAACCAGAGGGTAGTTTGACGGTTAAGTAGTCTAAGCA
TCCTTCACCGTTTCCACACTCCCAAGAGCTGAACTCTAAACCAGCAGCTCTCTGGAGCTACTGCTCTCC
CTCCACGTCGCCGTGTCCCTTGCCCTTCCCCTCAGGGCCGCAGACCGGCCGAGCCGCCGCAGCCGCCCG
CCGTTGGCCCGGCGTCCTGCGGGAAGCCGAGGGGGCCTCCCCGGGCCACCGCGCGAGCCGCTCCGCACC
ACAGGACGAGACAAACCGCGGCTATGTCGCCTTAGCCCTCGGGGTCCCACAGCCTCAGCAGCGTCCTAG
CCTGCCCGCTCCATGCCACGGCAAGGCTGCACCGTGTTCCAGGGGTGAAGGGGGCGATCGGGCATGCTC
CTCCCCATGGGTCGCCCACCATGTCTAATGGATATCGCACTCTGTCCCAGCACCTCAATGACCTGAAGA
AGGAGAACTTCAGCCTCAAGCTGCGCATCTACTTCCTGGAGGAGCGCATGCAACAGAAGTATGAAGTCA
GCCGGGAGGACGTCTACAAGCGGAACATTGAGCTGAAGGTTGAAGTGGAGAGCCTGAAACGAGAGCTCC
AGGACAGGAAACAGCATCTACATAAAACATGGGCCGATGAGGAGGATCTCAACAGCCAGAATGAAGCAG
AGCTCCGGCGCCAGGTTGAAGAACCGCAGCAGGAGACAGAACACGTTTATGAGCTCCTAGACAACAACA
TTCAGCTGCTGCAGGAGGAATCCAGGTTTGCAAAGGATGAAGCCACACAGATGGAGACTCTGGTGGAGG
CAGAGAAGGGGTGTAATCTGGAGCTCTCAGAGAGGTGGAAGGATGCTACCAAGAACAGGGAAGATGCAC
CGGGAGACCAGGTGAAGCTTGACCAATATTCTGCGGCACTGGCTCAGAGGGACAGGAGAATTGAAGAGC
TGAGGCAGAGCTTGGCTGCCCAGGAGGGGCTTGTGGAACAGCTGTCTCGAGAGAAACAACAACTGTTAC
ATCTGCTGGAGGAGCCTGGGGGCATGGAAGTGCAGCCCATGCCTAAAGGGTTACCCACGCAACAAAAGC
CAGACCTAAATGAGACCCCTACAACCCAGCCATCTGTGTCTGATTCCCACCTGGCAGAACTCCAGGACA
AAATCCAGCAAACAGAGGTCACCAACAAGATTCTTCAAGAGAAACTGAATGACATGAGCTGTGAGCTCA
GATCTGCACAGGAGTCGTCTCAGAAGCAAGATACGACAATCCAAAGCCTCAAGGAAATGCTAAAGAGCA
GGGAAAGTGAGACTGAAGAGCTGTACCAGGTGATTGAAGGTCAAAATGACACAATGGCAAAGCTTCCGG
AAATGCTACACCAGAGCCAGCTCGGACAGCTCCAGAGCTCAGAGGGCATTGCCCCTGCTCAGCAGCAAG
TGGCCCTGCTTGACCTTCAGAGTGCTCTGTTCTGCAGCCAGCTTGAAATCCAGAAGCTCCAGAGGCTGT
TACGCCAGAAAGAGCGTCAGCTGGCTGACGGCAAGCGGTGCATGCAATTTGTGGAGGCTGCAGCACAGG
AGAGAGAGCAGCAGAAGGAAGCTGCTTGGAAACATAACCAGGAATTACGAAAAGCTTTGCAACACCTCC
AAGGAGAACTGCACAGTAAGAGCCAACAGCTCCACGTTCTGGAGGCAGAAAAATATAATGAAATTCGAA
CCCAGGGACAAAACATTCAACACCTAAGTCACAGTCTGAGTCACAAAGAGCAGCTAATTCAGGAACTTC
AGGAGCTCCTACAGTATCGGGATACCACAGACAAAACTCTAGACACAAATGAGGTGTTTCTTGAGAAAC
TACGGCAACGAATACAAGACCGGGCAGTTGCTCTAGAGCGGGTTATAGATGAAAAGTTCTCTGCTCTAG
AAGAAAAGGACAAGGAACTGCGGCAGCTCCGGCTTGCTGTGAGGGACCGAGACCATGACTTAGAGAGAC
TGCGTTGTGTCCTGTCTGCCAATGAAGCTACCATGCAAAGTATGGAGAGTCTCCTGAGGGCCAGAGGCC
TGGAAGTGGAGCAGTTAATTGCCACCTGCCAAAACCTCCAGTGGTTGAAGGAAGAATTGGAAACCAAGT
TTGGCCACTGGCAGAAGGAACAGGAGAGCATCATTCAGCAGTTACAGACATCTCTGCATGACAGGAACA
AAGAAGTAGAGGATCTCAGTGCAACTTTGCTCCACAAACTTGGACCCGGCCAGAGTGAAGTAGCTGAGG
AGCTGTGCCAGCGCCTGCAGCGGAAGGAAAGGGTGCTGCAGGACCTTCTGAGTGATCGGAACAAACAAG
CCATGGAGCACGAGATGGAGGTCCAGGGACTGCTCCAGTCGATGGGCACCCGGGAACAGGAAAGACAGG
CTGTTGCAGAAAAAATGGTACAAGCCTTCATGGAAAGAAACTCGGAATTACAGGCCCTGCGGCAGTATC
TAGGGGGGAAGGAATTAATGGCAGCATCTCAGGCATTCATCTCTAACCAACCAGCTGGAGCGACTTCTG
TAGGCCCCCACCATGGAGAGCAAACTGACCAAGGTTCTACGCAGATGCCCTCTCGAGACGACAGCACCT
CGCTGACTGCCAGAGAGGAGGCCAGCATACCCCGGTCTACATTAGGAGACTCAGACACAGTTGCAGGGC
TGGAGAAAGAACTGAGCAATGCCAAGGAGGAGCTTGAGCTCATGGCCAAAAAAGAAAGAGAAAGCCAGA
TAGAATTGTCTGCCCTGCAGTCCATGATGGCTGTGCAAGAGGAAGAGCTGCAGGTGCAGGCTGCTGACT
TGGAGTCCCTGACCAGGAACATACAGATAAAAGAAGACCTCATAAAGGACCTGCAAATGCAACTGGTTG

FIG. 2 (CONT)

ACCCTGAAGATATGCCAGCCATGGAGCGCCTGACCCAAGAGGTCTTACTTCTTCGGGAAAAAGTTGCTT
CAGTGGAACCCCAGGGTCAGGAAGGGTCAGAGAACAGGAGACAACAGTTGCTGCTGATGTTAGAAGGAC
TAGTGGATGAACGGAGTCGGCTCAACGAGGCCCTGCAAGCTGAGCGGCAGCTCTACAGCAGCCTGGTCA
AGTTCCATGCCCAACCAGAGATCTCTGAGAGAGACCGAACTCTGCAGGTGGAACTGGAAGGGGCCCAGG
TGTTACGCAGTCGACTAGAAGAAGTTCTTGGAAGAAGCCTGGAGCGCTTAAGCAGGCTGGAGACCCTGG
CCGCCATTGGAGGTGCTACTGCAGGCGATGAGACTGAAGATACAAGCACACAGTTCACAGACAGCATTG
AGGAGGAGGCTGCACACAACAGCCACCAGCAACTCATCAAGGTGTCTTTGGAGAAAAGCCTGACCACCA
TGGAGACCCAGAACACATGTCTTCAGCCCCCTTCCCCAGTAGGAGAGGATGGTAACAGGCATCTTCAGG
AAGAAATGCTCCACCTGAGGGCTGAAATCCACCAGCCCTTAGAAGAGAAGAGAAAAGCTGAGGCAGAAC
TCAAGGAGCTAAAGGCTCAAATTGAGGAAGCAGGATTCTCCTCTGTGTCCCACATCAGGAACACCATGC
TGAGCCTTTGCCTTTGCCTTGAGAATGCAGAGCTGAAAGAGCAGATGGGAGAAGCAATGTCTGATGGAT
GGGAGGTGGAGGAAGACAAGGAGAAGGGCGAGGTGATGGTGGAGACCGTGGTGGCCAAAGGGGGTCTGA
GTGAGGACAGCCTTCAGGCTGAGTTCAGGAAAGTCCAGGGGAGACTCAAGAGTGCCTACAACATCATCA
ACCTCCTCAAAGAGCAGCTGGTCCTGAGAAGCTCGGAAGGGAACACTAAGGAGATGCCAGAGTTCCTCG
TGCGCCTGGCCAGGGAGGTGGACAGAATGAACATGGGCTTGCCTTCCTCGGAGAAGCATCAACACCAAG
AACAGGAGAATATGACCGCAAGGCCTGGCCCCAGGCCCCAGAGTCTCAAGCTTGGGACAGCTCTCTCAG
TAGACGGCTACCAACTGGAGAACAAGTCCCAGGCCCAAGACTCTGGACATCAGCCAGAATTTAGCCTAC
CAGGGTCCACCAAACACCTGCGCTCCCAGCTGGCTCAGTGTAGACAACGGTACCAAGATCTCCAGGAGA
AGCTGCTCATCTCAGAAGCCACTGTGTTTGCCCAGGCAAACCAGCTAGAGAAGTACAGAGCCATATTAA
GTGAATCCCTGGTGAAGCAGGACAGCAAGCAGATCCAGGTGGACCTTCAGGACCTGGGCTATGAGACTT
GTGGCCGAAGTGAGAATGAAGCTGAACGTGAGGAGACCACCAGCCCTGAGTGTGAGGAGCACGGTAACC
TGAAGCCTGTGGTGCTGGTGGAAGGCTTGTGCTCTGAGCAAGGGTACCTGGACCCTGTCTTGGTCAGCT
CACCTGTGAAGAACCCTTGGAGAACAAGCCAGGAAGCCAGAAGAATCCAGGCACAAGGAACTTCAGACA
ACAGCTCTCTCCTGAGGAAGGACATCCGAAATCTGAAAGCCCAGCTACCGAATGCCTACAAGGTCCTTC
AGAACCTGAGGAGCCGGGTCCGGTCCCTGTCTGCCACAAGCGATTACTCATCGAGTCTGGAGAGACCCC
GCAAGCTGATAGCCGTGGCAACCCTTGAGGGGGCCTCACCCCACAGTGTCACTGATGAAGACGAAGGCT
TGTTGTCAGATGGCACCGGGGCTTTTTACCCTCCAGGGCTCCAGGCCAAAAAGAATCTAGAGAATCTCA
TCCAGAGAGTATCCCAGCTGGAGGCCCAGCTCCCCAAAACTGGACTAGAAGGGAAGCTGGCTGAAGAAC
TGAAGTCCGCCTCGTGGCCTGGAAAATACGATTCTTTGATTCAGGATCAGGCCCGAAAAACTGTCATAT
CTGCGTCCGAAAATACXAAAAGGGAGAAGGATTTGTTTTCTTCTCACCCAACATTCGAAAGATACGTCA
AATCTTTTGAAGACCTCCTGAGGAACAACGACTTGACTACTTACCTGGGCCAGAGCTTCCGGGAACAAC
TTAGTTCAAGGCGTTCAGTGACAGACAGGCTGACCAGCAAATTCAGCACAAAGGATCATAAGAGTGAAA
AAGAAGAAGTTGGGCTTGAGCCACTGGCCTTCAGGTTCAGCAGGGAATTACAGGAGAAAGAGAAAGTGA
TTGAAGTCCTGCAGGCCAAGGTGGATACCCGGTTTTTCTCACCCCCCAGCAGCCATGCTGCGTCTGAGT
CCCACCGTTGTGCCAGCAGCACATCTTTCCTGTCGGATGACATAGAAGCCTGCTCTGACATGGACGTAG
CCAGCGAGTACACACACTATGAAGAGAAGAAGCCCTCACCCAGTAACTCAGCAGCCAGTGCATCTCAGG
GGCTTAAGGGCGAGCCCAGAAGCAGCTCCATCAGCTTGCCAACTCCCCAGAACCCCCCTAAGGAGGCCA
GCCAGGCCCAGCCAGGCTTTCACTTTAACTCCATACCCAAGCCGGCTAGCCTTTCCCAGGCACCAATGC
ACTTCACTGTACCCAGCTTCATGCCTTTCGGCCCCTCTGGGCCTCCCCTTCTTGGTTGCTGTGAGACAC
CAGTGGTGTCCTTGGCTGAGGCTCAACAAGAGCTGCAGATGCTGCAGAAGCAGCTGGGACGAAGTGTTA
GCATTGCCCCTCCCACCTCCACATCCACGTTGCTTAGCAACCACACAGAAGCTAGCTCTCCCCGCTACA
GCAACCCTGCTCAGCCCCACTCCCCAGCAAGGGGCACCATAGAGCTGGGCAGAATCCTGGAGCCTGGAT
ACCTGGGCAGCGGCCAGTGGGACATGATGAGGCCTCAGAAAGGGAGCATCTCTGGGGAGCTGTCCTCAG
GCTCCTCGATGTACCAGCTTAACTCCAAACCCACAGGGGCCGACCTGTTGGAAGAGCATTTAGGTGAGA
TCCGGAACCTGCGCCAGCGCCTGGAGGAGTCCATATGTGTCAATGACAGGCTACGGGAGCAGCTGCAGC
ATAGGCTCAGCTCCACGGCCCGAGAAAATGGTTCCACCTCTCACTTCTACAGTCAGGGCCTGGAGTCCA
TGCCTCAGCTCTACAATGAGAACAGAGCCCTCAGGGAAGAAAACCAAAGCCTGCAGACACGGCTCAGTC
ATGCTTCCAGGGGACACTCCCAGGAAGTGGACCACCTGAGGGAGGCTCTGCTTTCCTCAAGTTCCCAGC
TCCAGGAGCTGGAGAAGGAGCTGGAGCAGCAGAAGGCTGAGCGGCGGCAGCTTCTGGAAGACTTGCAGG
AGAAGCAGGATGAGATCGTGCATTTCCGAGAGGAGAGGCTGTCCCTCCAGGAAAACAACTCCAGGCTGC
AGCACAAGCTGGCCCTCCTGCAACAACAGTGTGAGGAGAAACAGCAGCTCTCCCTGTCCCTGCAGTCAG
AGCTCCAGATCTACGAGTCCCTCTACGAAAATCCTAAGAAGGGCTTGAAAGCCTTCAGCCTAGATTCCT
GTTACCAAGTCCCGGGTGAGTTGAGCTGCCTGGTGGCAGAGATTCGAGCTCTGAGAGTGCAGTTGGAGC
AGAGCATTCAAGTGAACAACCGTCTGCGGCTGCAGCTGGAACAGCAGATGGATCACGGTGCTGGCAAAG
CCAGTCTCAGTTCCTGCCCTGTTAACCAGAGCTTCTCAGCCAAGGCGGAGCTGGCAAACCAGCAGCCAC
CCTTCCAAGGTTCAGCTGCTTCCCCTCCAGTCCGGGACGTTGGCTTGAATTCTCCACCCGTGGTCCTCC
CCAGCAATTCGTGGTCTGTTCCTGGCTCAGACTCTGCCATCATCAGTAGGACAAACAATGGTTCGGATG
AGTCTGCAGCAACGAAGACCCCTCCCAAGATGGAGGTCGATGCTGCTGATGGCCCATTTGCCAGTGGAC
ACGGCAGACACGTCATCGGCCATGTGGATGACTACGACGCCCTACAGCAGCAGATTGGGGAAGGGAAGC
TGCTGATCCAAAAGATACTGTCTCTCACGAGGCCAGCACGCAGCGTCCCTGCACTGGACGCGCAGGGCA

FIG.2 (CONT)

CAGAGGCACCAGGTACCAAAAGTGTCCATGAGCTTCGGAGCAGCGCCAGGGCTCTGAACCACAGCCTAG
AAGAGTCAGCTTCCCTCCTCACCATGTTCTGGAGAGCAGCTTTGCCAAACTCTCATGGTTCTGTACTGG
TAGGCGAAGAGGGAAACCTGATGGAGAAAGAACTCCTAGACCTGCGAGCCCAAGTGTCCCAACAGCAAC
AGCTCCTTCAGAGCACTGCTGTGCGTCTGAAGACGGCCAACCAGAGGAAGAAAAGCATGGAGCAGTTCA
TCGTGAGCCATCTGACCAGGACCCATGATGTCTTGAAGAAAGCACGGACTAATTTAGAGATGAAATCCT
TCAGGGCCCTGATGTGCACTCCAGCCTTGTGACCCTTGCCTTCCAGGAGCCACATAAAAGGCGAAGCCA
GGAGTCCTTAAAACAGCAGGAAGGGTGGGCCTGCCCGCCCCTAGTACAGCTGCCTGTCTGCTGAGGAAT
ACCTGGTCCGACTCCTCCCCTGCTGGAGCTCCAGGGAAGGGCTCATATATGTGTCCACATGGGACAGGC
AGGAAGGAAAGTGGCATCCTGACAATGAATATGATTAGCCAAGGCCCACTGGGCCCATCACTAAGCAAA
ACTCATGTAGACTGTGTAGAAGGCCCCCCGGCACTGCTTCTAGACAGCCTCAGCAGCACGGTGCCCACC
TCGTTACAGTTCTCACCTCAAGATAGCCAACTCAGGGGAACTAGGACCTTACCACCCACAAACAGGATG
TGTGGTCCCAATGCCAACGCTCCTCAGACAGTTGTAAAAGCACACATCATTGAGTGGCAGCGTCCAGCC
GGACACTGTTGGAGACTACCAAACCCCTCACTGACCCAGTCTTGGGCCAGGCCAGCTCTGTGGGCCAAG
TCTGGTAGTACTTTGGTCTCTACCACCACACCAGAGAGAGTCTATATAGCAAATGTGGTAACTTGTAGG
TGCCCTGCACTTAGCCTAGCACCTTCTGTTTCTTACGTGATCTCAAGTTGAACCAACTTCCTTAACTCT
GCTGTCCCCTGAATCCTAACTTCCCTCAGGGGAATTGGAGATTGGTGGCCACATCATGCCTATTGAATG
TTTAGTGAACAGCATATCGGTGCCTCTTAATGGCATGGGCAAGGCCTGCTCTGTACTGAAGACTGTGTC
TTCACAGTGCTCATAGGACGTGGGTGTGTGTATAAATGTATAATATAGATTTATATATGTCGCTATGGC
TATGTGTTGAAGGCCAGCATAAGTGCAGAGCGATGGGTGAGAAGACGCTAAGCAGTCTTTCTTATGGCT
ATTAAAGCTAACTGTGTAC (SEQ ID NO:02)

FIG. 3.

MYOMEGALIN firstMET until stop

ATGTCTAATGGATATCGCACTCTGTCCCAGCACCTCAATGACCTGAAGAAGGAGAACTTCAGCCTCAAG
CTGCGCATCTACTTCCTGGAGGAGCGCATGCAACAGAAGTATGAAGTCAGCCGGGAGGACGTCTACAAG
CGGAACATTGAGCTGAAGGTTGAAGTGGAGAGCCTGAAACGAGAGCTCCAGGACAGGAAACAGCATCTA
CATAAAACATGGGCCGATGAGGAGGATCTCAACAGCCAGAATGAAGCAGAGCTCCGGCGCCAGGTTGAA
GAACCGCAGCAGGAGACAGAACACGTTTATGAGCTCCTAGACAACAACATTCAGCTGCTGCAGGAGGAA
TCCAGGTTTGCAAAGGATGAAGCCACACAGATGGAGACTCTGGTGGAGGCAGAGAAGGGGTGTAATCTG
GAGCTCTCAGAGAGGTGGAAGGATGCTACCAAGAACAGGGAAGATGCACCGGGAGACCAGGTGAAGCTT
GACCAATATTCTGCGGCACTGGCTCAGAGGGACAGGAGAATTGAAGAGCTGAGGCAGAGCTTGGCTGCC
CAGGAGGGGCTTGTGGAACAGCTGTCTCGAGAGAAACAACAACTGTTACATCTGCTGGAGGAGCCTGGG
GGCATGGAAGTGCAGCCCATGCCTAAAGGGTTACCCACGCAACAAAAGCCAGACCTAAATGAGACCCCT
ACAACCCAGCCATCTGTGTCTGATTCCCACCTGGCAGAACTCCAGGACAAAATCCAGCAAACAGAGGTC
ACCAACAAGATTCTTCAAGAGAAACTGAATGACATGAGCTGTGAGCTCAGATCTGCACAGGAGTCGTCT
CAGAAGCAAGATACGACAATCCAAAGCCTCAAGGAAATGCTAAAGAGCAGGGAAAGTGAGACTGAAGAG
CTGTACCAGGTGATTGAAGGTCAAAATGACACAATGGCAAAGCTTCCGGAAATGCTACACCAGAGCCAG
CTCGGACAGCTCCAGAGCTCAGAGGGCATTGCCCCTGCTCAGCAGCAAGTGGCCCTGCTTGACCTTCAG
AGTGCTCTGTTCTGCAGCCAGCTTGAAATCCAGAAGCTCCAGAGGCTGTTACGCCAGAAAGAGCGTCAG
CTGGCTGACGGCAAGCGGTGCATGCAATTTGTGGAGGCTGCAGCACAGGAGAGAGAGCAGCAGAAGGAA
GCTGCTTGGAAACATAACCAGGAATTACGAAAAGCTTTGCAACACCTCCAAGGAGAACTGCACAGTAAG
AGCCAACAGCTCCACGTTCTGGAGGCAGAAAAATATAATGAAATTCGAACCCAGGGACAAAACATTCAA
CACCTAAGTCACAGTCTGAGTCACAAAGAGCAGCTAATTCAGGAACTTCAGGAGCTCCTACAGTATCGG
GATACCACAGACAAAACTCTAGACACAAATGAGGTGTTTCTTGAGAAACTACGGCAACGAATACAAGAC
CGGGCAGTTGCTCTAGAGCGGGTTATAGATGAAAAGTTCTCTGCTCTAGAAGAAAAGGACAAGGAACTG
CGGCAGCTCCGGCTTGCTGTGAGGGACCGAGACCATGACTTAGAGAGACTGCGTTGTGTCCTGTCTGCC
AATGAAGCTACCATGCAAAGTATGGAGAGTCTCCTGAGGGCCAGAGGCCTGGAAGTGGAGCAGTTAATT
GCCACCTGCCAAAACCTCCAGTGGTTGAAGGAAGAATTGGAAACCAAGTTTGGCCACTGGCAGAAGGAA
CAGGAGAGCATCATTCAGCAGTTACAGACATCTCTGCATGACAGGAACAAAGAAGTAGAGGATCTCAGT
GCAACTTTGCTCCACAAACTTGGACCCGGCCAGAGTGAAGTAGCTGAGGAGCTGTGCCAGCGCCTGCAG
CGGAAGGAAAGGGTGCTGCAGGACCTTCTGAGTGATCGGAACAAACAAGCCATGGAGCACGAGATGGAG
GTCCAGGGACTGCTCCAGTCGATGGGCACCCGGGAACAGGAAAGACAGGCTGTTGCAGAAAAAATGGTA
CAAGCCTTCATGGAAAGAAACTCGGAATTACAGGCCCTGCGGCAGTATCTAGGGGGGAAGGAATTAATG
GCAGCATCTCAGGCATTCATCTCTAACCAACCAGCTGGAGCGACTTCTGTAGGCCCCCACCATGGAGAG
CAAACTGACCAAGGTTCTACGCAGATGCCCTCTCGAGACGACAGCACCTCGCTGACTGCCAGAGAGGAG
GCCAGCATACCCCGGTCTACATTAGGAGACTCAGACACAGTTGCAGGGCTGGAGAAAGAACTGAGCAAT
GCCAAGGAGGAGCTTGAGCTCATGGCCAAAAAAGAAAGAGAAAGCCAGATAGAATTGTCTGCCCTGCAG
TCCATGATGGCTGTGCAAGAGGAAGAGCTGCAGGTGCAGGCTGCTGACTTGGAGTCCCTGACCAGGAAC
ATACAGATAAAAGAAGACCTCATAAAGGACCTGCAAATGCAACTGGTTGACCCTGAAGATATGCCAGCC
ATGGAGCGCCTGACCCAAGAGGTCTTACTTCTTCGGGAAAAAGTTGCTTCAGTGGAACCCCAGGGTCAG
GAAGGGTCAGAGAACAGGAGACAACAGTTGCTGCTGATGTTAGAAGGACTAGTGGATGAACGGAGTCGG
CTCAACGAGGCCCTGCAAGCTGAGCGGCAGCTCTACAGCAGCCTGGTCAAGTTCCATGCCCAACCAGAG
ATCTCTGAGAGAGACCGAACTCTGCAGGTGGAACTGGAAGGGGCCCAGGTGTTACGCAGTCGACTAGAA
GAAGTTCTTGGAAGAAGCCTGGAGCGCTTAAGCAGGCTGGAGACCCTGGCCGCCATTGGAGGTGCTACT
GCAGGCGATGAGACTGAAGATACAAGCACACAGTTCACAGACAGCATTGAGGAGGAGGCTGCACACAAC
AGCCACCAGCAACTCATCAAGGTGTCTTTGGAGAAAAGCCTGACCACCATGGAGACCCAGAACACATGT
CTTCAGCCCCCTTCCCCAGTAGGAGAGGATGGTAACAGGCATCTTCAGGAAGAAATGCTCCACCTGAGG
GCTGAAATCCACCAGCCCTTAGAAGAGAAGAGAAAAGCTGAGGCAGAACTCAAGGAGCTAAAGGCTCAA
ATTGAGGAAGCAGGATTCTCCTCTGTGTCCCACATCAGGAACACCATGCTGAGCCTTTGCCTTTGCCTT
GAGAATGCAGAGCTGAAAGAGCAGATGGGAGAAGCAATGTCTGATGGATGGGAGGTGGAGGAAGACAAG
GAGAAGGGCGAGGTGATGGTGGAGACCGTGGTGGCCAAAGGGGGTCTGAGTGAGGACAGCCTTCAGGCT
GAGTTCAGGAAAGTCCAGGGGAGACTCAAGAGTGCCTACAACATCATCAACCTCCTCAAAGAGCAGCTG
GTCCTGAGAAGCTCGGAAGGGAACACTAAGGAGATGCCAGAGTTCCTCGTGCGCCTGGCCAGGGAGGTG
GACAGAATGAACATGGGCTTGCCTTCCTCGGAGAAGCATCAACACCAAGAACAGGAGAATATGACCGCA
AGGCCTGGCCCCAGGCCCCAGAGTCTCAAGCTTGGGACAGCTCTCTCAGTAGACGGCTACCAACTGGAG
AACAAGTCCCAGGCCCAAGACTCTGGACATCAGCCAGAATTTAGCCTACCAGGGTCCACCAAACACCTG
CGCTCCCAGCTGGCTCAGTGTAGACAACGGTACCAAGATCTCCAGGAGAAGCTGCTCATCTCAGAAGCC
ACTGTGTTTGCCCAGGCAAACCAGCTAGAGAAGTACAGAGCCATATTAAGTGAATCCCTGGTGAAGCAG
GACAGCAAGCAGATCCAGGTGGACCTTCAGGACCTGGGCTATGAGACTTGTGGCCGAAGTGAGAATGAA
GCTGAACGTGAGGAGACCACCAGCCCTGAGTGTGAGGAGCACGGTAACCTGAAGCCTGTGGTGCTGGTG
GAAGGCTTGTGCTCTGAGCAAGGGTACCTGGACCCTGTCTTGGTCAGCTCACCTGTGAAGAACCCTTGG
AGAACAAGCCAGGAAGCCAGAAGAATCCAGGCACAAGGAACTTCAGACAACAGCTCTCTCCTGAGGAAG
GACATCCGAAATCTGAAAGCCCAGCTACCGAATGCCTACAAGGTCCTTCAGAACCTGAGGAGCCGGGTC

FIG. 3 (cont)

CGGTCCCTGTCTGCCACAAGCGATTACTCATCGAGTCTGGAGAGACCCCGCAAGCTGATAGCCGTGGCA
ACCCTTGAGGGGGCCTCACCCCACAGTGTCACTGATGAAGACGAAGGCTTGTTGTCAGATGGCACCGGG
GCTTTTTACCCTCCAGGGCTCCAGGCCAAAAAGAATCTAGAGAATCTCATCCAGAGAGTATCCCAGCTG
GAGGCCCAGCTCCCCAAAACTGGACTAGAAGGGAAGCTGGCTGAAGAACTGAAGTCCGCCTCGTGGCCT
GGAAAATACGATTCTTTGATTCAGGATCAGGCCCGAAAAACTGTCATATCTGCGTCCGAAAATACXAAA
AGGGAGAAGGATTTGTTTTCTTCTCACCCAACATTCGAAAGATACGTCAAATCTTTTGAAGACCTCCTG
AGGAACAACGACTTGACTACTTACCTGGGCCAGAGCTTCCGGGAACAACTTAGTTCAAGGCGTTCAGTG
ACAGACAGGCTGACCAGCAAATTCAGCACAAAGGATCATAAGAGTGAAAAAGAAGAAGTTGGGCTTGAG
CCACTGGCCTTCAGGTTCAGCAGGGAATTACAGGAGAAAGAGAAAGTGATTGAAGTCCTGCAGGCCAAG
GTGGATACCCGGTTTTTCTCACCCCCCAGCAGCCATGCTGCGTCTGAGTCCCACCGTTGTGCCAGCAGC
ACATCTTTCCTGTCGGATGACATAGAAGCCTGCTCTGACATGGACGTAGCCAGCGAGTACACACACTAT
GAAGAGAAGAAGCCCTCACCCAGTAACTCAGCAGCCAGTGCATCTCAGGGGCTTAAGGGCGAGCCCAGA
AGCAGCTCCATCAGCTTGCCAACTCCCCAGAACCCCCCTAAGGAGGCCAGCCAGGCCCAGCCAGGCTTT
CACTTTAACTCCATACCCAAGCCGGCTAGCCTTTCCCAGGCACCAATGCACTTCACTGTACCCAGCTTC
ATGCCTTTCGGCCCCTCTGGGCCTCCCCTTCTTGGTTGCTGTGAGACACCAGTGGTGTCCTTGGCTGAG
GCTCAACAAGAGCTGCAGATGCTGCAGAAGCAGCTGGGACGAAGTGTTAGCATTGCCCCTCCCACCTCC
ACATCCACGTTGCTTAGCAACCACACAGAAGCTAGCTCTCCCCGCTACAGCAACCCTGCTCAGCCCCAC
TCCCCAGCAAGGGGCACCATAGAGCTGGGCAGAATCCTGGAGCCTGGATACCTGGGCAGCGGCCAGTGG
GACATGATGAGGCCTCAGAAAGGGAGCATCTCTGGGGAGCTGTCCTCAGGCTCCTCGATGTACCAGCTT
AACTCCAAACCCACAGGGGCCGACCTGTTGGAAGAGCATTTAGGTGAGATCCGGAACCTGCGCCAGCGC
CTGGAGGAGTCCATATGTGTCAATGACAGGCTACGGGAGCAGCTGCAGCATAGGCTCAGCTCCACGGCC
CGAGAAAATGGTTCCACCTCTCACTTCTACAGTCAGGGCCTGGAGTCCATGCCTCAGCTCTACAATGAG
AACAGAGCCCTCAGGGAAGAAAACCAAAGCCTGCAGACACGGCTCAGTCATGCTTCCAGGGGACACTCC
CAGGAAGTGGACCACCTGAGGGAGGCTCTGCTTTCCTCAAGTTCCCAGCTCCAGGAGCTGGAGAAGGAG
CTGGAGCAGCAGAAGGCTGAGCGGCGGCAGCTTCTGGAAGACTTGCAGGAGAAGCAGGATGAGATCGTG
CATTTCCGAGAGGAGAGGCTGTCCCTCCAGGAAAACAACTCCAGGCTGCAGCACAAGCTGGCCCTCCTG
CAACAACAGTGTGAGGAGAAACAGCAGCTCTCCCTGTCCCTGCAGTCAGAGCTCCAGATCTACGAGTCC
CTCTACGAAAATCCTAAGAAGGGCTTGAAAGCCTTCAGCCTAGATTCCTGTTACCAAGTCCCGGGTGAG
TTGAGCTGCCTGGTGGCAGAGATTCGAGCTCTGAGAGTGCAGTTGGAGCAGAGCATTCAAGTGAACAAC
CGTCTGCGGCTGCAGCTGGAACAGCAGATGGATCACGGTGCTGGCAAAGCCAGTCTCAGTTCCTGCCCT
GTTAACCAGAGCTTCTCAGCCAAGGCGGAGCTGGCAAACCAGCAGCCACCCTTCCAAGGTTCAGCTGCT
TCCCCTCCAGTCCGGGACGTTGGCTTGAATTCTCCACCCGTGGTCCTCCCCAGCAATTCGTGCTCTGTT
CCTGGCTCAGACTCTGCCATCATCAGTAGGACAAACAATGGTTCGGATGAGTCTGCAGCAACGAAGACC
CCTCCCAAGATGGAGGTCGATGCTGCTGATGGCCCATTTGCCAGTGGACACGGCAGACACGTCATCGGC
CATGTGGATGACTACGACGCCCTACAGCAGCAGATTGGGGAAGGGAAGCTGCTGATCCAAAAGATACTG
TCTCTCACGAGGCCAGCACGCAGCGTCCCTGCACTGGACGCGCAGGGCACAGAGGCACCAGGTACCAAA
AGTGTCCATGAGCTTCGGAGCAGCGCCAGGGCTCTGAACCACAGCCTAGAAGAGTCAGCTTCCCTCCTC
ACCATGTTCTGGAGAGCAGCTTTGCCAAACTCTCATGGTTCTGTACTGGTAGGCGAAGAGGGAAACCTG
ATGGAGAAAGAACTCCTAGACCTGCGAGCCCAAGTGTCCCAACAGCAACAGCTCCTTCAGAGCACTGCT
GTGCGTCTGAAGACGGCCAACCAGAGGAAGAAAAGCATGGAGCAGTTCATCGTGAGCCATCTGACCAGG
ACCCATGATGTCTTGAAGAAAGCACGGACTAATTTAGAGATGAAATCCTTCAGGGCCCTGATGTGCACT
CCAGCCTTGTGA (SEQ ID NO:03)

FIG. 4

Human mycmegalin cDNA

```
   1  GGATCCTTGA GGGCACTGGT GCGACTTTCA GGTGAGGTCT TAGCAGATGA
  51  AAGCGGCTGG CTGTGGCCCG CGCCAGTAGT GCTTTCTGCT CCGCACTCGC
 101  CGTGAGCCAG GTGTGCAACC GGATTTGGGG CGAGGGTCGC GCTGGCTACC
 151  TCGCATGCGC AGAGCCGGAA GCCCGCTGAC CGGACTACAG CTCCCAGAAG
 201  AGCCTTGTGG AGGCCGCAGA CGCGAAGCCG CTGGCGCCAT CTTGAAATCT
 251  GATCCTCCAT CCCCGAGGCT TTGCGTCTGC GCGGCCGGCC GCTGCTGCTC
 301  CGGGAGCCCA GTCTGCTAAA AGGGGAGGAC GTTGAGGACG CGGCGGCTGG
 351  CGGGAGAGAC AGCTGGGGAG AGACATGGCA GGGTCGGAGC GCGGCCTGCG
 401  CCTCTGTCAC TCAGCATCCT CTTAGGCGTT TCCACGCCCG CCCCCTGCCC
 451  GAGGGGCGGG GCTGACGGCT CTGGTACCCG GAGTCGGCGC GCGGGGCAGG
 501  GGCGCGCCCC TGCAGAGTGG GGACCCCACT GGGCTGTGCC ATGCTGACCG
 551  GAGACCACCG AGGCGGGAGA CAGAGCGCGG CGAAGAGCCA TTGAGTGGTC
 601  ACCCAGTAGC CGCCGCCGCC GCCGCCTCGG GAAGCTTGCC ACCCGCTAGG
 651  AGGGAAGATG AAGGAGATTT GCAGGATCTG TGCCCGAGAG CTGTGTGGAA
 701  ACCAGCGGCG CTGGATCTTC CACACGGCGT CCAAGCTCAA TCTCCAGGTT
 751  CTGCTTTCGC ACGTCTTGGG CAAGGATGTC CCCCGCGATG GCAAAGCCGA
 801  GTTCGCTTGC AGCAAGTGTG CTTTCATGCT TGATCGAATC TATCGATTCG
 851  ACACAGTTAT TGCCCGGATT GAAGCGCTTT CTATTGAGCG CTTGCAAAAG
 901  CTGCTACTGG AGAAGGATCG CCTCAAGTTC TGCATTGCCA GTATGTATCG
 951  GAAGAATAAC GATGACTCTG GCGCGGAGAT CAAGGCGGGG AATGGGACGG
1001  TTGACATGTC CGTCTTACCC GATGCGAGAT ACTCTGCACT GCTCCAGGAG
1051  GACTTCGCCT ATTCAGGGTT TGAGTGCTGG GTGGAGAATG AGGATCAGAT
1101  CCAGGAGCCA CACAGCTGCC ATGGTTCAGA AGGCCCTGGA AACCGACCCA
1151  GGAGATGCCG TGGTTGTGCC GCTTTGCGGG TTGCTGATTC TGACTATGAA
1201  GCCATTTGTA AGGTACCTCG AAAGGTGGCC AGAAGTATCT CCTGCGGCCC
1251  TTCTAGCAGG TGGTCGACCA GCATTTGCAC TGAAGAACCA GCGTTGTCTG
1301  AGGTTGGGCC ACCCGACTTA GCAAGCACAA AGGTACCCCC AGATGGAGAA
1351  AGCATGGAGG AAGAGACGCC TGGTTCCTCT GTGGAATCTT TGGATGCAAG
1401  CGTCCAGGCT AGCCCTCCAC AACAGAAAGA TGAGGAGACT GAGAGAAGTG
1451  CAAAGGAACT TGGAAAGTGT GACTGTTGTT CAGATGATCA GGCTCCGCAG
1501  CATGGGTGTA ATCACAAGCT GGAATTAGCT CTTAGCATGA TTAAAGGTCT
1551  TGATTATAAG CCCATCCAGA GCCCCCGAGG GAGCAGGCTT CCGATTCCAG
1601  TGAAATCCAG CCTACCTGGA GCCAAGCCTG GCCCTAGCAT GACAGATGGA
1651  GTTAGTTCCG GTTTCCTTAA CAGGTCTTTG AAACCCCTTT ACAAGACACC
1701  TGTGAGTTAT CCCTTGGAGC TTTCAGACCT GCAGGAGCTG TGGGATGATC
1751  TCTGTGAAGA TTATTTGCCG CTCCGGGTCC AGCCCATGAC TGAAGAGTTG
1801  CTGAAACAAC AAAAGCTGAA TTCACATGAG ACCACTATAA CTCAGCAGTC
1851  TGTATCTGAT TCCCACTTGG CAGAACTCCA GGAAAAAATC CAGCAAACAG
1901  AGGCCACCAA CAAGATTCTT CAAGAGAAAC TTAATGAAAT GAGCTATGAA
1951  CTAAAGTGTG CTCAGGAGTC GTCTCAAAAG CAAGATGGTA CAATTCAGAA
2001  CCTCAAGGAA ACTCTGAAAA GCAGGGAACG TGAGACTGAG GAGTTGTACC
2051  AGGTAATTGA AGGTCAAAAT GACACAATGG CAAAGCTTCG AGAAATGCTG
2101  CACCAAAGCC AGCTTGGACA ACTTCACAGC TCAGAGGGTA CTTCTCCAGC
2151  TCAGCAACAG GTAGCTCTGC TTGATCTTCA GAGTGCTTTA TTCTGCAGCC
2201  AACTTGAAAT ACAGAAGCTC CAGAGGGTGG TACGACAGAA AGAGCGCCAA
2251  CTGGCTGATG CCAAACAATG TGTGCAATTT GTAGAGGCTG CAGCACACGA
2301  GAGTGAACAG CAGAAAGAGG CTTCTTGGAA ACATAACCAG GAATTGCGAA
2351  AAGCCTTGCA GCAGCTACAA GAAGAATTGC AGAATAAGAG CCAACAGCTT
2401  CGTGCCTGGG AGGCTGAAAA ATACAATGAG ATTCGAACCC AGGAACAAAA
2451  CATCCAGCAC CTAAACCATA GTCTGAGTCA CAAGGAGCAG TTGCTTCAGG
2501  AATTTCGGGA GCTCCTACAG TATCGAGATA ACTCAGACAA AACCCTTGAA
2551  GCAAATGAAA TGTTGCTTGA GAAACTTCGC CAGCGAATAC ATGATAAAGC
2601  TGTTGCTCTG GAGCGGGCTA TAGATGAAAA ATTCTCTGCT CTAGAAGAGA
2651  AAGAAAAAGA ACTGCGCCAG CTTCGTCTTG CTGTGAGAGA GCGAGATCAT
2701  GACTTAGAGA GACTGCGCGA TGTCCTCTCC TCCAATGAAG CTACTATGCA
2751  AAGTATGGAG AGTCTCCTGA GGGCCAAAGG CCTGGAAGTG GAACAGTTAT
2801  CTACTACCTG TCAAAACCTC CAGTGGCTGA AAGAAGAAAT GGAAACCAAA
2851  TTTAGGCGTT GGCAGAAGGA ACAAGAGAGT ATCATTCAGC AGTTACAGAC
2901  GTCTCTTCAT GATAGGAACA AAGAAGTGGA GGATCTTAGT GCAACACTGC
```

FIGURE 4 (CONT)

```
2951  TCTGCAAACT TGGACCAGGG CAGAGTGAGA TAGCAGAGGA GCTGTGCCAG
3001  CGTCTACAGC GAAAGGAAAG GATGCTGCAG GACCTTCTAA GTGATCGAAA
3051  TAAACAAGTG CTGGAACATG AAATGGAGAT TCAAGGCCTG CTTCAGTCTG
3101  TGAGCACCAG GGAGCAGGAA AGCCAAGCTG CTGCAGAGAA GTTGGTGCAA
3151  GCCTTAATGG AAAGAAATTC AGAATTACAG GCCCTGCGCC AATATTTAGG
3201  AGGGAGAGAC TCCCTGATGT CCCAAGCACC CATCTCTAAC CAACAAGCTG
3251  AAGTTACCCC CACTGGCCGT CTTGGAAAAC AGACTGATCA AGGTTCAATG
3301  CAGATACCTT CCAGAGATGA TAGCACTTCA TTGACTGCCA AAGAGGATGT
3351  CAGCATACCC AGATCCACAT TAGGAGACTT GGACACAGTT GCAGGGCTGG
3401  AAAAAGAACT GAGTAATGCC AAAGAGGAAC TTGAACTCAT GGCTAAAAAA
3451  GAAAGAGAAA GTCAGATGGA ACTTTCTGCT CTACAGTCCA TGATGGCTGT
3501  GCAGGAAGAA GAGCTGCAGG TGCAGGCTGC TGATATGGAG TCTCTGACCA
3551  GGAACATACA GATTAAAGAA GATCTCATAA AGGACCTGCA AATGCAACTG
3601  GTTGATCCTG AAGACATACC AGCTATGGAA CGCCTGACCC AGGAAGTCTT
3651  ACTTCTTCGG GAAAAAGTTG CTTCAGTAGA ATCCCAGGGT CAAGAAATTT
3701  CAGGAAACCG AAGACAACAG TTGCTGCTGA TGCTAGAAGG ACTAGTAGAT
3751  GAACGGAGTC GGCTCAATGA GGCCTTACAA GCAGAGAGAC AGCTCTATAG
3801  CAGTCTGGTG AAGTTCCATG CCCATCCAGA GAGCTCTGAG AGAGACCGAA
3851  CTCTGCAGGT GGAACTGGAA GGGGCTCAGG TGTTACGCAG TCGGCTAGAA
3901  GAAGTTCTTG GAAGAAGCTT GGAGCGCTTA AACAGGCTGG AGACCCTGGC
3951  CGCCATTGGA GGTGCAGCTG CAGGGGATGA CACCGAAGAT ACAAGCACTG
4001  AGTTCACTGA CAGTATTGAG GAGGAGGCTG CACACCATAG TCACCAGCAA
4051  CTTGTCAAGG TGGCTTTGGA GAAAAGTCTG GCAACTGTGG AGACCCAGAA
4101  CCCATCTTTT TCCCCTCCTT CTCCGATGGG AGGGGACAGT AACAGGTGTC
4151  TTCAGGAAGA AATGCTCCAC CTGAGGGCTG AGTTCCACCA GCACTTAGAA
4201  GAGAAGAGGA AAGCTGAGGA GGAACTGAAG GAGCTAAAGG CTCAAATTGA
4251  GGAAGCAGGA TTCTCCTCAG TGTCCCACAT CAGGAACACC ATGCTGAGCC
4301  TTTGCCTTGA GAATGCGGAG CTGAAAGAGC AGATGGGAGA AGCAATGTCT
4351  GATGGATGGG AGATCGAGGA AGACAAGGAG AAGGGCGAGG TGATGGTTGA
4401  GACTGTGGTA ACCAAAGAGG GTCTGAGTGA GAGTAGCCTT CAGGCTGAGT
4451  TCAGAAAGCT CCAGGGAAAA CTGAAGAATG CCCACAATAT CATCAACCTC
4501  CTCAAAGAAC AACTTGTGCT GAGTAGCAAG GAAGGGAATA GTAAACTTAC
4551  TCCAGAGCTC CTTGTGCATC TGACCAGCAC CATTGAAAGA ATAAACACAG
4601  AACTGGTTGG TTCCCCTGGG AAGCACCAAC ACCAAGAGGA GGGGAATGTG
4651  ACTGTGAGGC CTTTCCCCAG ACCCCAGAGC CTTGACCTTG GGGCTACCTT
4701  CACAGTGGAT GCCCACCAAT TGGATAACCA GTCCCAGCCT CGTGACCCTG
4751  GGCCTCAGTC AGCGTTTAGC CTACCAGGGT CCACCCAGCA CCTGCGCTCC
4801  CAGCTGTCAC AATGCAAACA ACGCTATCAA GATCTCCAGG AGAAGCTGCT
4851  GCTATCAGAA GCCACTGTCT TTGCTCAGGC TAACGAGCTG GAGAAATACA
4901  GAGTTATGCT TACAGGTGAA TCCTTGGTGA AGCAGGACAG CAAGCAGATC
4951  CAGGTGGACC TCCAGGACCT GGGCTATGAG ACTTGTGGCC GAAGCGAGAA
5001  TGAGGCTGAA CGGGAGGAAA CCACCAGTCC TGAGTGTGAG GAGCACAACA
5051  GCCTCAAGGA AATGGTCCTG ATGGAGGGGC TGTGCTCTGA GCAGGGACGC
5101  CGGGGCTCAA CACTGGCTAG TTCCTCTGAG AGGAAGCCCT TGGAGAACCA
5151  GCTAGGGAAG CAGGAAGAGT TCCGGGTATA TGGAAAGTCA GAAAACATCT
5201  TGGTCCTACG AAAGGACATC AAAGATCTGA AGGCCCAGCT GCAGAATGCC
5251  AACAAGGTCA TTCAAAACCT CAAGAGCCGG GTCCGGTCCC TCTCAGTTAC
5301  AAGTGATTAT TCGTCTAGTC TGGAAAGACC CCGGAAGCTG AGAGCTGTTG
5351  GCACCTTGGA GGGGTCTTCA CCTCATAGTG TCCCTGATGA GGATGAGGGG
5401  TGGCTGTCTG ATGGCACTGG GGCTTTCTAC TCTCCAGGGC TTCAGGCCAA
5451  AAAGGACCTG GAGAGTCTCA TCCAGAGAGT ATCCCAGCTG GAGGCCCAGC
5501  TCCCAAAAAA TGGACTAGAA GAGAAGCTGG CTGAGGAGCT GAGATCAGCC
5551  TCGTGGCCTG GGAAATATGA TTCCCTGATT CAGGATCAGG CCCGGGAACT
5601  GTCTTACCTA CGGCAAAAAA TACGAGAAGG GAGAGGTATT TGTTATCTTA
5651  TCACCCGGCA TGCAAAAGAT ACAGTAAAAT CTTTTGAGGA TCTCCTAAGG
5701  AGCAATGACA TTGACTACTA CCTGGGACAG AGCTTCCGGG AGCAACTCGC
5751  CCAGGGAAGC CAGCTGACAG AGAGGCTCAC CAGCAAACTC AGCACCAAGG
5801  ATCATAAAAG TGAGAAAGAT CAAGCTGGAC TTGAGCCACT GGCCCTCAGG
5851  CTCAGCAGGG AGCTGCAGGA GAAGGAGAAA GTGATTGAAG TCCTGCAGGC
5901  CAAGCTGGAT GCTCGGTCCC TCACACCCTC CAGCAGCCAT GCCTTGTCTG
5951  ACTCCCACCG CTCTCCCAGC AGCACCTCTT TCCTGTCTGA TGAACTGGAA
```

FIGURE 4(CONT)

```
6001  GCCTGCTCTG ACATGGACAT AGTCAGCGAG TACACACACT ATGAAGAGAA
6051  GAAAGCTTCT CCCAGTCACT CAGATTCCAT CCATCATTCG AGTCATTCTG
6101  CTGTGTTGTC TTCTAAACCA TCATCAACCA GTGCATCTCA GGGGGCTAAG
6151  GCCGAATCCA ACAGCAACCC CATCAGCTTG CCAACTCCCC AGAATACCCC
6201  CAAGGAGGCC AACCAGGCCC ATTCAGGCTT TCATTTTCAC TCCATACCCA
6251  AGCTGGCTAG CCTTCCTCAG GCACCATTGC CCTCAGCTCC ATCCAGCTTC
6301  CTGCCTTTCA GCCCCACTGG CCCTCTCCTC CTTGGCTGCT GTGAGACACC
6351  AGTGGTCTCC TTGGCTGAGG CTCAGCAGGA GCTACAGATG CTGCAGAAGC
6401  AGTTGGGAGA AAGTGCCAGC ACTGTTCCTC CTGCTTCCAC AGCTACATTG
6451  CTGAGCAACG ACTTGGAAGC CGACTCTTCC TACTACCTCA ACTCTGCCCA
6501  GCCTCACTCT CCTCCAAGGG GCACCATAGA ACTGGGAAGA ATCCTAGAGC
6551  CTGGGTACCT GGGCAGCAGT GGCAAGTGGG ATGTGATGAG GCCTCAGAAA
6601  GGGAGTGTAT CTGGGGACCT ATCCTCAGGC TCCTCTGTGT ACCAGCTTAA
6651  CTCCAAACCC ACAGGGGCTG ACCTGCTGGA AGAGCATCTT GGTGAAATCC
6701  GGAACCTGCG CCAGCGCCTG GAGGAGTCCA TCTGCATCAA TGACCGCCTA
6751  CGGGAGCAAC TGGAACACCG GCTGACCTCT ACTGCTCGTG GAAGGGGATC
6801  CACTTCTAAC TTCTACAGTC AGGGCCTGGA GTCCATACCT CAGCTCTGCA
6851  ATGAGAACAG AGTCCTCAGG GAAGACAATC GAAGACTTCA GGCTCAACTG
6901  AGTCATGTTT CCAGAGAGCA CTCCCAGGAA ACAGAAAGCC TGAGGGAGGC
6951  TCTGCTGTCC TCTCGATCCC ACCTTCAAGA GCTGGAAAAG GAGCTGGAGC
7001  ACCAGAAGGT GGAAAGGCAG CAGCTTTTGG AAGACTTGAG GGAGAAGCAG
7051  CAAGAGGTCT TGCATTTCAG GGAGGAACGT CTTTCCCTCC AGGAAAACGA
7101  CTCCAGTGGG CCTTGCCTCT CCCTGGTCAG ACTGCAGCAC AAGCTGGTTC
7151  TCCTGCAGCA ACAGTGTGAA GAGAAACAGC AGCTCTTTGA GTCCCTCCAG
7201  TCAGAGCTAC AAATCTACGA GGCACTTTAT GGCAATTCCA AGAAGGGGCT
7251  GAAAGCTTAC AGCCTGGATG CCTGTCACCA AATCCCTTTG AGCAGTGACC
7301  TGAGCCACCT GGTGGCAGAG GTACGAGCTC TGAGAGGGCA GCTGGAGCAG
7351  AGCATTCAGG GGAACAATTG TCTGCGACTG CAGCTGCAAC AGCAGCTGGA
7401  GAGCGGTGCT GGCAAAGCCA GCCTCAGCCC CTCCTCCATT AACCAGAACT
7451  TCCCAGCCAG CACTGACCCT GGAAACAAGC AGCTGCTCCT CCAAGATTCA
7501  GCTGTGTCCC CTCCAGTCCG GGATGTTGGT ATGAATTCCC CAGCTCTGGT
7551  CTTCCCCAGC TCTGCTTCCT CTACTCCTGG CTCAGAAACG CCCATAATCA
7601  ACAGAGCAAA TGGCTTGGGT TTGGATACTT CTCCAGTAAT GAAGACCCCT
7651  CCCAAGCTAG AGGGTGATGC TACTGATGGC TCCTTTGCCA ATAAGCATGG
7701  CCGCCATGTC ATTGGCCACA TTGATGACTA CAGTGCCCTA AGACAGCAGA
7751  TTGCGGAGGG CAAGCTGCTG GTCAAAAAGA TAGTGTCTCT TGTGAGATCA
7801  GCGTGCAGCT TCCCTGGCCT TGAAGCCCAA GGCACAGAGG TGCTAGGCAG
7851  CAAAGGTATT CATGAGCTTC GGAGCAGCAC CAGTGCCCTG CACCATGCCC
7901  TAGAGGAGTC GGCTTCCCTC CTCACCATGT TCTGGAGAGC AGCCCTGCCA
7951  AGCACCCACA TCCCTGTGCT GCCTGGCAAA GTGGGAGAAT CAACAGAAAG
8001  GGAACTTCTG GAACTGAGAA CCAAAGTATC CAAACAGGAG CGGCTCCTTC
8051  AGAGCACAAC TGAGCATCTG AAGAACGCCA ACCAGCAGAA GGAGAGCATG
8101  GAGCAGTTCA TCGTCAGCCA GCTAACCAGA ACACATGATG TTTTAAAGAA
8151  GGCAAGGACT AACTTAGAGG TGAAATCCCT AAGGGCTCTG CCATGTACTC
8201  CAGCCTTGTG ACCCTTGCCT TCCAGGAACC ATGCAAGAAG CGCAGCCACC
8251  AGAAGTCCTT AAAACAGCAG GAAAGGTGGG CCTGTCCCCC TTTTGTGCAG
8301  CTACCTATCT GCTGAGGAGC ATCTGGGCCT CATTCCTCCA AGTCCACGGG
8351  AGGGTCCAGA AGAGGGAGTC AGAGATGTAT CCTGGTGGAG CTGGGAGAAA
8401  GGCAGAAAGC CTTTCTGACA GCTATGGAAT ACGATTAGCC AAGGTCCACT
8451  TGGCCCAGCA CTAAGAAAAA GATGCGTAGT TTGCACAGAA GGTTTTGTGA
8501  TCCTGCCTCT CAACAGCCCC AGCAGCTTGG GAACTAGCAA GAGCACATTT
8551  CTTGCCTCAT CAGCTGTCCT GAGATGGAAA ACTCAGTGGA TATAGGACCC
8601  TGATTCCGAT GAAAGGGGCA CGTGGTCCCA ATGCTGGAGC TCCTCTGGCA
8651  GGTTCTAAAA GCACACTACT GAGCAGCGGT GCCCTGCCGG ACACTGCTGG
8701  CGGGGGCTCA GTGAGCACTA CTCACAGATC CACACCTGAC CCTGTTGGGT
8751  CGAGTCAGGC TGGGCCTTGG TCTGCACTGT AGCACCTGTG TTCTTTGAGT
8801  TCACATCATG AATGTGGTGA CTTCCCAGAT ACCATCTCAG GCTTAACCTA
8851  GCACATCCTA TTTCTTTTCT TCTATGATAT CCAAATTGGA CTGACCTCAC
8901  TTCAAAGTTG CTGTCCCATT TTGTCACCCT ATCTTATCTC GGGGAAATTG
8951  CAGACTGATG GCCAGACCAA CTCTGTTGAA ATTCTTGCAT AGAGCAAACC
9001  TGTGCTCATT TTTAAGTGGC ATGGGAGAGG CCCCCAGCCT AGTAAAGCCT
```

FIGURE 4 (CONT)

```
9051  AGTCTGTGTC TTCACAGTGC TGGTAGAATG TGTTTGTGTG TATAAATATA
9101  TGATATAGAT TTATATATGT TGCTAACGCC ATATATTGAA GGCCAACATA
9151  ACTGGTGGAC AGGGTGGGTG ACAGAAAATG AAAGCCTTTT TGGTGATTGT
9201  TAAAGCAAGA TGTGTATAAA GAAATAAATA GTTTTTCTTT C (SEQ ID NO:04)
```

FIG. 5

```
>Human myomegalin protein
   1  MKEICRICAR ELCGNQRRWI FHTASKLNLQ VLLSHVLGKD VPRDGKAEFA
  51  CSKCAFMLDR IYRFDTVIAR IEALSIERLQ KLLLEKDRLK FCIASMYRKN
 101  NDDSGAEIKA GNGTVDMSVL PDARYSALLQ EDFAYSGFEC WVENEDQIQE
 151  PHSCHGSEGP GNRPRRCRGC AALRVADSDY EAICKVPRKV ARSISCGPSS
 201  RWSTSICTEE PALSEVGPPD LASTKVPPDG ESMEEETPGS SVESLDASVQ
 251  ASPPQQKDEE TERSAKELGK CDCCSDDQAP QHGCNHKLEL ALSMIKGLDY
 301  KPIQSPRGSR LPIPVKSSLP GAKPGPSMTD GVSSGFLNRS LKPLYKTPVS
 351  YPLELSDLQE LWDDLCEDYL PLRVQPMTEE LLKQQKLNSH ETTITQQSVS
 401  DSHLAELQEK IQQTEATNKI LQEKLNEMSY ELKCAQESSQ KQDGTIQNLK
 451  ETLKSRERET EELYQVIEGQ NDTMAKLREM LHQSQLGQLH SSEGTSPAQQ
 501  QVALLDLQSA LFCSQLEIQK LQRVVRQKER QLADAKQCVQ FVEAAAHESE
 551  QQKEASWKHN QELRKALQQL QEELQNKSQQ LRAWEAEKYN EIRTQEQNIQ
 601  HLNHSLSHKE QLLQEFRELL QYRDNSDKTL EANEMLLEKL RQRIHDKAVA
 651  LERAIDEKFS ALEEKEKELR QLRLAVRERD HDLERLRDVL SSNEATMQSM
 701  ESLLRAKGLE VEQLSTTCQN LQWLKEEMET KFSRWQKEQE SIIQQLQTSL
 751  HDRNKEVEDL SATLLCKLGP GQSEIAEELC QRLQRKERML QDLLSDRNKQ
 801  VLEHEMEIQG LLQSVSTREQ ESQAAAEKLV QALMERNSEL QALRQYLGGR
 851  DSLMSQAPIS NQQAEVTPTG RLGKQTDQGS MQIPSRDDST SLTAKEDVSI
 901  PRSTLGDLDT VAGLEKELSN AKEELELMAK KERESQMELS ALQSMMAVQE
 951  EELQVQAADM ESLTRNIQIK EDLIKDLQMQ LVDPEDIPAM ERLTQEVLLL
1001  REKVASVESQ GQEISGNRRQ QLLLMLEGLV DERSRLNEAL QAERQLYSSL
1051  VKFHAHPESS ERDRTLQVEL EGAQVLRSRL EEVLGRSLER LNRLETLAAI
1101  GGAAAGDDTE DTSTEFTDSI EEEAAHHSHQ QLVKVALEKS LATVETQNPS
1151  FSPPSPMGGD SNRCLQEEML HLRAEFHQHL EEKRKAEEEL KELKAQIEEA
1201  GFSSVSHIRN TMLSLCLENA ELKEQMGEAM SDGWEIEEDK EKGEVMVETV
1251  VTKEGLSESS LQAEFRKLQG KLKNAHNIIN LLKEQLVLSS KEGNSKLTPE
1301  LLVHLTSTIE RINTELVGSP GKHQHQEEGN VTVRPFPRPQ SLDLGATFTV
1351  DAHQLDNQSQ PRDPGPQSAF SLPGSTQHLR SQLSQCKQRY QDLQEKLLLS
1401  EATVFAQANE LEKYRVMLTG ESLVKQDSKQ IQVDLQDLGY ETCGRSENEA
1451  EREETTSPEC EEHNSLKEMV LMEGLCSEQG RRGSTLASSS ERKPLENQLG
1501  KQEEFRVYGK SENILVLRKD IKDLKAQLQN ANKVIQNLKS RVRSLSVTSD
1551  YSSSLERPRK LRAVGTLEGS SPHSVPDEDE GWLSDGTGAF YSPGLQAKKD
1601  LESLIQRVSQ LEAQLPKNGL EEKLAEELRS ASWPGKYDSL IQDQARELSY
1651  LRQKIREGRG ICYLITRHAK DTVKSFEDLL RSNDIDYYLG QSFREQLAQG
1701  SQLTERLTSK LSTKDHKSEK DQAGLEPLAL RLSRELQEKE KVIEVLQAKL
1751  DARSLTPSSS HALSDSHRSP SSTSFLSDEL EACSDMDIVS EYTHYEEKKA
1801  SPSHSDSIHH SSHSAVLSSK PSSTSASQGA KAESNSNPIS LPTPQNTPKE
1851  ANQAHSGFHF HSIPKLASLP QAPLPSAPSS FLPFSPTGPL LLGCCETPVV
1901  SLAEAQQELQ MLQKQLGESA STVPPASTAT LLSNDLEADS SYYLNSAQPH
1951  SPPRGTIELG RILEPGYLGS SGKWDVMRPQ KGSVSGDLSS GSSVYQLNSK
2001  PTGADLLEEH LGEIRNLRQR LEESICINDR LREQLEHRLT STARGRGSTS
2051  NFYSQGLESI PQLCNENRVL REDNRRLQAQ LSHVSREHSQ ETESLREALL
2101  SSRSHLQELE KELEHQKVER QQLLEDLREK QQEVLHFREE RLSLQENDSS
2151  GPCLSLVRLQ HKLVLLQQQC EEKQQLFESL QSELQIYEAL YGNSKKGLKA
2201  YSLDACHQIP LSSDLSHLVA EVRALRGQLE QSIQGNNCLR LQLQQQLESG
2251  AGKASLSPSS INQNFPASTD PGNKQLLLQD SAVSPPVRDV GMNSPALVFP
2301  SSASSTPGSE TPIINRANGL GLDTSPVMKT PPKLEGDATD GSFANKHGRH
2351  VIGHIDDYSA LRQQIAEGKL LVKKIVSLVR SACSFPGLEA QGTEVLGSKG
2401  IHELRSSTSA LHHALEESAS LLTMFWRAAL PSTHIPVLPG KVGESTEREL
2451  LELRTKVSKQ ERLLQSTTEH LKNANQQKES MEQFIVSQLT RTHDVLKKAR
2501  TNLEVKSLRA LPCTPAL (SEQ ID NO:05)
```

FIGURE 6

M14 PROTEIN

MMAQFPTAMNGGPNMWAITSEERTKHDKQFDNLKPSGGYITGDQARTFFLQSGLPAPVL
AEIWALSDLNKDGKMDQQEFSIAMKLIKLKLQGQQLPVVLPPIMKQPPMFSPLISARFG
MGSMPNLSIHQPLPPVAPITAPLSSATSGTSIPPLMMPAPLVPSVSTSSLPNGTASLIQ
PLSIPYSSSTLPHASSYSLMMGGFGGASIQKAQSLIDLGSSSSTSSTASLSGNSPKTGT
SEWAVPQPSRLKYRQKFNSLDKSMSGYLSGFQARNALLQSNLSQTQLATIWTLADIDGD
GQLKAEEFILAMHLTDMAKAGQPLPLTLPPELVPPSFRGGKQIDSINGTLPSYQKTQEE
EPQKKLPVTFEDKRKANYERGNMELEKRRQVLMEQQQREAERKAQKEKEEWERKQRELQ
EQEWKKQLELEKRLEKQRELERQREEERRKEIERREAAKQELERQRRLEWERIRRQELL
NQKNREQEEIVRLNSKKKSLHLELEAVNGKHQQISGRLQDVRIRKQTQKTELEVLDKQC
DLEIMEIKQLQQELQEYQNKLIYLVPEKQLLNERIKNMQLSNTPDSGISLLHKKSSEKE
ELCQRLKEQLDALEKETASKLSEMDSFNNQLKCGNMDDSVLQCLLSLLSCLNNLFLLLK
ELRESYNTQQLALEQLHKIKRDKLKELERKRLEQIQKKKLEDEAARKAKQGKENLWKES
IRKEEEEKQKRLQEEKSQDRTQEEERKTEAKQSETARALVNYRALYPFEARNHDEMSFN
SGDIIQVDEKTVGEPGWLYGSFQGKFGWFPCNYVEKMLSSDKTPSPKKALLPPAVSLSA
TSAAPQPLCSNQPAPVTDYQNVSFSNLNVNTTWQQKSAFTRTVSPGSVSPIHGQGQAVE
NLKAQALCSWTAKKENHLNFSKHDVITVLEQQENWWFGEVHGGRGWFPKSYVKIIPGSE
VKRGEPEALYAAVNKKPTSTAYPVGEEYIALYSYSSVEPGDLTFTEGEELLVTQKDGEW
WTGSIGERTGIFPSNYVRPKDQENVGNASKSGASNKKPEIAQVTSAYAASGAEQLSLAP
GQLILILKKNSSGWWQGELQARGKKRQKGWFPASHVKLLGPSAERTTPAFHAVCQVIAM
YDYIANNEDELNFSKGQLINVMNKDDPDWWQGEINGVTGLFPSNYVKMTTDSDPSQQWC
ADLQALDTMQPMERKRQGYIHELIETEERYMDDLQLVIEVFQKRMAESGFLTEAEMALI
FVNWKELIMSNTKLLKALRVRKKTGGEKMPVEMMGDILAAELSHMQAYIRFCSCQLNGA
ALLQQKTDEDADFKEFLKKLASDPRCKGMPLSSFLLKPMQRITRYPLLIRSILENTPQN
HVDHSSLKLALERAEELCSQVNEGVREKENSDRLEWIQAHVQCEGLAEQLIFNSLTNCL
GPRKLLYSGKLYKTKSNKELHGFLFNDFLLLTYLVRQFAASSGFEKLFSSKSSAQFKMY
KTPIFLNEVLVKLPTDPSSDEPVFHISHIDRVYTLRTDNINERTAWVQKIKAASEQYID
TEKKKREKAYQARSQKTSGIGRLMVHVIEATELKACKPNGKSNPYCEISMGSQSYTTRT
LQDTLNPKWNFNCQFFIKDLYQDVLCLTMFDRDQFSPDDFLGRTEVPVAKIRTEQESKG
PTTRRLLLHEVPTGEVWVRFDLQLFEQKTLL (SEQ ID NO:08)

SEQUENCE LISTING

<110> Conti, Marco
Pahlke, Gudrun

<120> Novel Phosphodiesterase Interacting Proteins

<130> SUN-101PCT

<140> 60/108,255
<141> 1998-11-12

<160> 8

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2326
<212> PRT
<213> rat

<400> 1

```
Met Ser Asn Gly Tyr Arg Thr Leu Ser Gln His Leu Asn Asp Leu Lys
 1               5                  10                  15
Lys Glu Asn Phe Ser Leu Lys Leu Arg Ile Tyr Phe Leu Glu Glu Arg
            20                  25                  30
Met Gln Gln Lys Tyr Glu Val Ser Arg Glu Asp Val Tyr Lys Arg Asn
        35                  40                  45
Ile Glu Leu Lys Val Glu Val Glu Ser Leu Lys Arg Glu Leu Gln Asp
    50                  55                  60
Arg Lys Gln His Leu His Lys Thr Trp Ala Asp Glu Glu Asp Leu Asn
65                  70                  75                  80
Ser Gln Asn Glu Ala Glu Leu Arg Arg Gln Val Glu Glu Pro Gln Gln
                85                  90                  95
Glu Thr Glu His Val Tyr Glu Leu Leu Asp Asn Asn Ile Gln Leu Leu
            100                 105                 110
Gln Glu Glu Ser Arg Phe Ala Lys Asp Glu Ala Thr Gln Met Glu Thr
        115                 120                 125
Leu Val Glu Ala Glu Lys Gly Cys Asn Leu Glu Leu Ser Glu Arg Trp
    130                 135                 140
Lys Asp Ala Thr Lys Asn Arg Glu Asp Ala Pro Gly Asp Gln Val Lys
145                 150                 155                 160
Leu Asp Gln Tyr Ser Ala Ala Leu Ala Gln Arg Asp Arg Arg Ile Glu
                165                 170                 175
Glu Leu Arg Gln Ser Leu Ala Ala Gln Glu Gly Leu Val Glu Gln Leu
            180                 185                 190
Ser Arg Glu Lys Gln Gln Leu Leu His Leu Leu Glu Glu Pro Gly Gly
        195                 200                 205
Met Glu Val Gln Pro Met Pro Lys Gly Leu Pro Thr Gln Gln Lys Pro
    210                 215                 220
Asp Leu Asn Glu Thr Pro Thr Thr Gln Pro Ser Val Ser Asp Ser His
225                 230                 235                 240
Leu Ala Glu Leu Gln Asp Lys Ile Gln Gln Thr Glu Val Thr Asn Lys
                245                 250                 255
Ile Leu Gln Glu Lys Leu Asn Asp Met Ser Cys Glu Leu Arg Ser Ala
            260                 265                 270
Gln Glu Ser Ser Gln Lys Gln Asp Thr Thr Ile Gln Ser Leu Lys Glu
        275                 280                 285
Met Leu Lys Ser Arg Glu Ser Glu Thr Glu Glu Leu Tyr Gln Val Ile
    290                 295                 300
Glu Gly Gln Asn Asp Thr Met Ala Lys Leu Pro Glu Met Leu His Gln
```

```
305                 310                 315                 320
Ser Gln Leu Gly Gln Leu Gln Ser Ser Glu Gly Ile Ala Pro Ala Gln
                325                 330                 335
Gln Gln Val Ala Leu Leu Asp Leu Gln Ser Ala Leu Phe Cys Ser Gln
            340                 345                 350
Leu Glu Ile Gln Lys Leu Gln Arg Leu Leu Arg Gln Lys Glu Arg Gln
        355                 360                 365
Leu Ala Asp Gly Lys Arg Cys Met Gln Phe Val Glu Ala Ala Ala Gln
    370                 375                 380
Glu Arg Glu Gln Gln Lys Glu Ala Ala Trp Lys His Asn Gln Glu Leu
385                 390                 395                 400
Arg Lys Ala Leu Gln His Leu Gln Gly Glu Leu His Ser Lys Ser Gln
                405                 410                 415
Gln Leu His Val Leu Glu Ala Glu Lys Tyr Asn Glu Ile Arg Thr Gln
            420                 425                 430
Gly Gln Asn Ile Gln His Leu Ser His Ser Leu Ser His Lys Glu Gln
        435                 440                 445
Leu Ile Gln Glu Leu Gln Glu Leu Leu Gln Tyr Arg Asp Thr Thr Asp
    450                 455                 460
Lys Thr Leu Asp Thr Asn Glu Val Phe Leu Glu Lys Leu Arg Gln Arg
465                 470                 475                 480
Ile Gln Asp Arg Ala Val Ala Leu Glu Arg Val Ile Asp Glu Lys Phe
                485                 490                 495
Ser Ala Leu Glu Glu Lys Asp Lys Glu Leu Arg Gln Leu Arg Leu Ala
            500                 505                 510
Val Arg Asp Arg Asp His Asp Leu Glu Arg Leu Arg Cys Val Leu Ser
        515                 520                 525
Ala Asn Glu Ala Thr Met Gln Ser Met Glu Ser Leu Leu Arg Ala Arg
    530                 535                 540
Gly Leu Glu Val Glu Gln Leu Ile Ala Thr Cys Gln Asn Leu Gln Trp
545                 550                 555                 560
Leu Lys Glu Glu Leu Glu Thr Lys Phe Gly His Trp Gln Lys Glu Gln
                565                 570                 575
Glu Ser Ile Ile Gln Gln Leu Gln Thr Ser Leu His Asp Arg Asn Lys
            580                 585                 590
Glu Val Glu Asp Leu Ser Ala Thr Leu Leu His Lys Leu Gly Pro Gly
        595                 600                 605
Gln Ser Glu Val Ala Glu Glu Leu Cys Gln Arg Leu Gln Arg Lys Glu
    610                 615                 620
Arg Val Leu Gln Asp Leu Leu Ser Asp Arg Asn Lys Gln Ala Met Glu
625                 630                 635                 640
His Glu Met Glu Val Gln Gly Leu Leu Gln Ser Met Gly Thr Arg Glu
                645                 650                 655
Gln Glu Arg Gln Ala Val Ala Glu Lys Met Val Gln Ala Phe Met Glu
            660                 665                 670
Arg Asn Ser Glu Leu Gln Ala Leu Arg Gln Tyr Leu Gly Gly Lys Glu
        675                 680                 685
Leu Met Ala Ala Ser Gln Ala Phe Ile Ser Asn Gln Pro Ala Gly Ala
    690                 695                 700
Thr Ser Val Gly Pro His His Gly Glu Gln Thr Asp Gln Gly Ser Thr
705                 710                 715                 720
Gln Met Pro Ser Arg Asp Asp Ser Thr Ser Leu Thr Ala Arg Glu Glu
                725                 730                 735
Ala Ser Ile Pro Arg Ser Thr Leu Gly Asp Ser Asp Thr Val Ala Gly
            740                 745                 750
Leu Glu Lys Glu Leu Ser Asn Ala Lys Glu Glu Leu Glu Leu Met Ala
        755                 760                 765
Lys Lys Glu Arg Glu Ser Gln Ile Glu Leu Ser Ala Leu Gln Ser Met
    770                 775                 780
Met Ala Val Gln Glu Glu Glu Leu Gln Val Gln Ala Ala Asp Leu Glu
785                 790                 795                 800
Ser Leu Thr Arg Asn Ile Gln Ile Lys Glu Asp Leu Ile Lys Asp Leu
                805                 810                 815
```

```
Gln Met Gln Leu Val Asp Pro Glu Asp Met Pro Ala Met Glu Arg Leu
            820                 825                 830
Thr Gln Glu Val Leu Leu Leu Arg Glu Lys Val Ala Ser Val Glu Pro
        835                 840                 845
Gln Gly Gln Glu Gly Ser Glu Asn Arg Arg Gln Gln Leu Leu Leu Met
    850                 855                 860
Leu Glu Gly Leu Val Asp Glu Arg Ser Arg Leu Asn Glu Ala Leu Gln
865                 870                 875                 880
Ala Glu Arg Gln Leu Tyr Ser Ser Leu Val Lys Phe His Ala Gln Pro
                885                 890                 895
Glu Ile Ser Glu Arg Asp Arg Thr Leu Gln Val Glu Leu Glu Gly Ala
            900                 905                 910
Gln Val Leu Arg Ser Arg Leu Glu Glu Val Leu Gly Arg Ser Leu Glu
        915                 920                 925
Arg Leu Ser Arg Leu Glu Thr Leu Ala Ala Ile Gly Gly Ala Thr Ala
    930                 935                 940
Gly Asp Glu Thr Glu Asp Thr Ser Thr Gln Phe Thr Asp Ser Ile Glu
945                 950                 955                 960
Glu Glu Ala Ala His Asn Ser His Gln Gln Leu Ile Lys Val Ser Leu
                965                 970                 975
Glu Lys Ser Leu Thr Thr Met Glu Thr Gln Asn Thr Cys Leu Gln Pro
            980                 985                 990
Pro Ser Pro Val Gly Glu Asp Gly Asn Arg His Leu Gln Glu Glu Met
        995                 1000                1005
Leu His Leu Arg Ala Glu Ile His Gln Pro Leu Glu Glu Lys Arg Lys
    1010                1015                1020
Ala Glu Ala Glu Leu Lys Glu Leu Lys Ala Gln Ile Glu Glu Ala Gly
1025                1030                1035                1040
Phe Ser Ser Val Ser His Ile Arg Asn Thr Met Leu Ser Leu Cys Leu
                1045                1050                1055
Cys Leu Glu Asn Ala Glu Leu Lys Glu Gln Met Gly Glu Ala Met Ser
            1060                1065                1070
Asp Gly Trp Glu Val Glu Glu Asp Lys Glu Lys Gly Glu Val Met Val
        1075                1080                1085
Glu Thr Val Val Ala Lys Gly Gly Leu Ser Glu Asp Ser Leu Gln Ala
    1090                1095                1100
Glu Phe Arg Lys Val Gln Gly Arg Leu Lys Ser Ala Tyr Asn Ile Ile
1105                1110                1115                1120
Asn Leu Leu Lys Glu Gln Leu Val Leu Arg Ser Ser Glu Gly Asn Thr
                1125                1130                1135
Lys Glu Met Pro Glu Phe Leu Val Arg Leu Ala Arg Glu Val Asp Arg
            1140                1145                1150
Met Asn Met Gly Leu Pro Ser Ser Glu Lys His Gln His Gln Glu Gln
        1155                1160                1165
Glu Asn Met Thr Ala Arg Pro Gly Pro Arg Pro Gln Ser Leu Lys Leu
    1170                1175                1180
Gly Thr Ala Leu Ser Val Asp Gly Tyr Gln Leu Glu Asn Lys Ser Gln
1185                1190                1195                1200
Ala Gln Asp Ser Gly His Gln Pro Glu Phe Ser Leu Pro Gly Ser Thr
                1205                1210                1215
Lys His Leu Arg Ser Gln Leu Ala Gln Cys Arg Gln Arg Tyr Gln Asp
            1220                1225                1230
Leu Gln Glu Lys Leu Leu Ile Ser Glu Ala Thr Val Phe Ala Gln Ala
        1235                1240                1245
Asn Gln Leu Glu Lys Tyr Arg Ala Ile Leu Ser Glu Ser Leu Val Lys
    1250                1255                1260
Gln Asp Ser Lys Gln Ile Gln Val Asp Leu Gln Asp Leu Gly Tyr Glu
1265                1270                1275                1280
Thr Cys Gly Arg Ser Glu Asn Glu Ala Glu Arg Glu Glu Thr Thr Ser
                1285                1290                1295
Pro Glu Cys Glu Glu His Gly Asn Leu Lys Pro Val Val Leu Val Glu
            1300                1305                1310
Gly Leu Cys Ser Glu Gln Gly Tyr Leu Asp Pro Val Leu Val Ser Ser
```

1315 1320 1325
Pro Val Lys Asn Pro Trp Arg Thr Ser Gln Glu Ala Arg Arg Ile Gln
1330 1335 1340
Ala Gln Gly Thr Ser Asp Asn Ser Ser Leu Leu Arg Lys Asp Ile Arg
1345 1350 1355 1360
Asn Leu Lys Ala Gln Leu Pro Asn Ala Tyr Lys Val Leu Gln Asn Leu
1365 1370 1375
Arg Ser Arg Val Arg Ser Leu Ser Ala Thr Ser Asp Tyr Ser Ser Ser
1380 1385 1390
Leu Glu Arg Pro Arg Lys Leu Ile Ala Val Ala Thr Leu Glu Gly Ala
1395 1400 1405
Ser Pro His Ser Val Thr Asp Glu Asp Glu Gly Leu Leu Ser Asp Gly
1410 1415 1420
Thr Gly Ala Phe Tyr Pro Pro Gly Leu Gln Ala Lys Lys Asn Leu Glu
1425 1430 1435 1440
Asn Leu Ile Gln Arg Val Ser Gln Leu Glu Ala Gln Leu Pro Lys Thr
1445 1450 1455
Gly Leu Glu Gly Lys Leu Ala Glu Glu Leu Lys Ser Ala Ser Trp Pro
1460 1465 1470
Gly Lys Tyr Asp Ser Leu Ile Gln Asp Gln Ala Arg Lys Thr Val Ile
1475 1480 1485
Ser Ala Ser Glu Asn Thr Lys Arg Glu Lys Asp Leu Phe Ser Ser His
1490 1495 1500
Pro Thr Phe Glu Arg Tyr Val Lys Ser Phe Glu Asp Leu Leu Arg Asn
1505 1510 1515 1520
Asn Asp Leu Thr Thr Tyr Leu Gly Gln Ser Phe Arg Glu Gln Leu Ser
1525 1530 1535
Ser Arg Arg Ser Val Thr Asp Arg Leu Thr Ser Lys Phe Ser Thr Lys
1540 1545 1550
Asp His Lys Ser Glu Lys Glu Glu Val Gly Leu Glu Pro Leu Ala Phe
1555 1560 1565
Arg Phe Ser Arg Glu Leu Gln Glu Lys Glu Lys Val Ile Glu Val Leu
1570 1575 1580
Gln Ala Lys Val Asp Thr Arg Phe Phe Ser Pro Pro Ser Ser His Ala
1585 1590 1595 1600
Ala Ser Glu Ser His Arg Cys Ala Ser Ser Thr Ser Phe Leu Ser Asp
1605 1610 1615
Asp Ile Glu Ala Cys Ser Asp Met Asp Val Ala Ser Glu Tyr Thr His
1620 1625 1630
Tyr Glu Glu Lys Lys Pro Ser Pro Ser Asn Ser Ala Ala Ser Ala Ser
1635 1640 1645
Gln Gly Leu Lys Gly Glu Pro Arg Ser Ser Ser Ile Ser Leu Pro Thr
1650 1655 1660
Pro Gln Asn Pro Pro Lys Glu Ala Ser Gln Ala Gln Pro Gly Phe His
1665 1670 1675 1680
Phe Asn Ser Ile Pro Lys Pro Ala Ser Leu Ser Gln Ala Pro Met His
1685 1690 1695
Phe Thr Val Pro Ser Phe Met Pro Phe Gly Pro Ser Gly Pro Pro Leu
1700 1705 1710
Leu Gly Cys Cys Glu Thr Pro Val Val Ser Leu Ala Glu Ala Gln Gln
1715 1720 1725
Glu Leu Gln Met Leu Gln Lys Gln Leu Gly Arg Ser Val Ser Ile Ala
1730 1735 1740
Pro Pro Thr Ser Thr Ser Thr Leu Leu Ser Asn His Thr Glu Ala Ser
1745 1750 1755 1760
Ser Pro Arg Tyr Ser Asn Pro Ala Gln Pro His Ser Pro Ala Arg Gly
1765 1770 1775
Thr Ile Glu Leu Gly Arg Ile Leu Glu Pro Gly Tyr Leu Gly Ser Gly
1780 1785 1790
Gln Trp Asp Met Met Arg Pro Gln Lys Gly Ser Ile Ser Gly Glu Leu
1795 1800 1805
Ser Ser Gly Ser Ser Met Tyr Gln Leu Asn Ser Lys Pro Thr Gly Ala
1810 1815 1820

Asp Leu Leu Glu Glu His Leu Gly Glu Ile Arg Asn Leu Arg Gln Arg
1825 1830 1835 1840
Leu Glu Glu Ser Ile Cys Val Asn Asp Arg Leu Arg Glu Gln Leu Gln
1845 1850 1855
His Arg Leu Ser Ser Thr Ala Arg Glu Asn Gly Ser Thr Ser His Phe
1860 1865 1870
Tyr Ser Gln Gly Leu Glu Ser Met Pro Gln Leu Tyr Asn Glu Asn Arg
1875 1880 1885
Ala Leu Arg Glu Glu Asn Gln Ser Leu Gln Thr Arg Leu Ser His Ala
1890 1895 1900
Ser Arg Gly His Ser Gln Glu Val Asp His Leu Arg Glu Ala Leu Leu
1905 1910 1915 1920
Ser Ser Ser Ser Gln Leu Gln Glu Leu Glu Lys Glu Leu Glu Gln Gln
1925 1930 1935
Lys Ala Glu Arg Arg Gln Leu Leu Glu Asp Leu Gln Glu Lys Gln Asp
1940 1945 1950
Glu Ile Val His Phe Arg Glu Glu Arg Leu Ser Leu Gln Glu Asn Asn
1955 1960 1965
Ser Arg Leu Gln His Lys Leu Ala Leu Leu Gln Gln Gln Cys Glu Glu
1970 1975 1980
Lys Gln Gln Leu Ser Leu Ser Leu Gln Ser Glu Leu Gln Ile Tyr Glu
1985 1990 1995 2000
Ser Leu Tyr Glu Asn Pro Lys Lys Gly Leu Lys Ala Phe Ser Leu Asp
2005 2010 2015
Ser Cys Tyr Gln Val Pro Gly Glu Leu Ser Cys Leu Val Ala Glu Ile
2020 2025 2030
Arg Ala Leu Arg Val Gln Leu Glu Gln Ser Ile Gln Val Asn Asn Arg
2035 2040 2045
Leu Arg Leu Gln Leu Glu Gln Gln Met Asp His Gly Ala Gly Lys Ala
2050 2055 2060
Ser Leu Ser Ser Cys Pro Val Asn Gln Ser Phe Ser Ala Lys Ala Glu
2065 2070 2075 2080
Leu Ala Asn Gln Gln Pro Pro Phe Gln Gly Ser Ala Ala Ser Pro Pro
2085 2090 2095
Val Arg Asp Val Gly Leu Asn Ser Pro Pro Val Val Leu Pro Ser Asn
2100 2105 2110
Ser Cys Ser Val Pro Gly Ser Asp Ser Ala Ile Ile Ser Arg Thr Asn
2115 2120 2125
Asn Gly Ser Asp Glu Ser Ala Ala Thr Lys Thr Pro Pro Lys Met Glu
2130 2135 2140
Val Asp Ala Ala Asp Gly Pro Phe Ala Ser Gly His Gly Arg His Val
2145 2150 2155 2160
Ile Gly His Val Asp Asp Tyr Asp Ala Leu Gln Gln Gln Ile Gly Glu
2165 2170 2175
Gly Lys Leu Leu Ile Gln Lys Ile Leu Ser Leu Thr Arg Pro Ala Arg
2180 2185 2190
Ser Val Pro Ala Leu Asp Ala Gln Gly Thr Glu Ala Pro Gly Thr Lys
2195 2200 2205
Ser Val His Glu Leu Arg Ser Ser Ala Arg Ala Leu Asn His Ser Leu
2210 2215 2220
Glu Glu Ser Ala Ser Leu Leu Thr Met Phe Trp Arg Ala Ala Leu Pro
2225 2230 2235 2240
Asn Ser His Gly Ser Val Leu Val Gly Glu Glu Gly Asn Leu Met Glu
2245 2250 2255
Lys Glu Leu Leu Asp Leu Arg Ala Gln Val Ser Gln Gln Gln Gln Leu
2260 2265 2270
Leu Gln Ser Thr Ala Val Arg Leu Lys Thr Ala Asn Gln Arg Lys Lys
2275 2280 2285
Ser Met Glu Gln Phe Ile Val Ser His Leu Thr Arg Thr His Asp Val
2290 2295 2300
Leu Lys Lys Ala Arg Thr Asn Leu Glu Met Lys Ser Phe Arg Ala Leu
2305 2310 2315 2320
Met Cys Thr Pro Ala Leu

2325

<210> 2
<211> 9679
<212> DNA
<213> rat

<220>
<221> misc_feature
<222> (1)...(9679)
<223> n = A,T,C or G

<400> 2
ccggtcccct ttggtagtag tatctcagag ctcgccccat agtttcatag ttcatgtctg 60
gtttgttctt atgctttccc cagagcttcg agacagcctt tgagtccacc agcttgaata 120
tgccctttc tctctgagtc catttaatat acctgggaca agtattttta tcttgaagca 180
gatctaaaag aaactcccac agataggttg tgtttccttt tccttctctg gctttcttct 240
tgactcctaa ctcaggagac ccattggaaa ctggtgactg ctgggtcttt ggtttacggc 300
caactttctt ctttttcatt ggttcgtggc tgtctggtga agtatggata ggcgaggcat 360
ccattggttc agactcctct gttgacacct ccactacagt ctccgtaatg acatctggcc 420
tcatcgcagc atggataaaa tcggattctt gaatcctcaa gcaggtagga gactccatat 480
gaagcagggc ttcagcagct tcaatggtct tatccgtaca gtgtgcatta ctgctgtgaa 540
ctgatgcttc cacggcccgg atgagcagat ccagctggtt cgtgggtccc tcatgcagag 600
acgtcgccat gtttatcccg gggctggaac tgctgcttca cattgactta caccctgagc 660
agcggcgaca ggggagaagg cggaacccgc ggccggagac acacgccgtg cgggcggcac 720
acactcacgc actcgcacac actccgacgc ccggatcctt gcgcgtcctc cgacaggaag 780
cggcggccgg cccgcctccc gccgccgggc tgagcagccc caccacctaa cggcaggggc 840
ggcggcggcc ccggctggca acgcgatcct tccgccccgc gcccagacag gaagtcccgg 900
gcgccggcag ccagcggccg cacggacacc tgaggctggg gagcccgcag gccgccctcg 960
gggacgcggg cctcggcagg aaaaggcgcg cttcacgttc tgcggaagcg aagtctgcaa 1020
atgtcccctc agcatggtct tcctcctggc tcaatctgtc tcaccttcag gtgatcctag 1080
gactggggct cctttccagg tccccagttt ctcaagtcga tcttctacct ccctcttgat 1140
tttctactcc attgctggaa agctccagaa cagagcctcc gccgccaacc actgctgatg 1200
ccatcgcgtc ttccctgagc aagtttcgaa cgctgcgaat caatgtaatt acggctcaga 1260
tgattgccag ggttatcggt ttcatgttct aattcaatag tgatggagta gacatccaga 1320
agtccagtct tctaaagatg attaaccaga gggtagtttg acggttaagt agtctaagca 1380
tccttcaccg tttccacact cccaagagct gaactctaaa ccagcagctc tctggagcta 1440
ctgctctccc tccacgtcgc cgtgtccctt gcccttcccc tcagggccgc agaccggccg 1500
agccgccgca gccgcccgcc gttggcccgg cgtcctgcgg gaagccgagg gggcctcccc 1560
gggccaccgc gcgagccgct ccgcaccaca ggacgagaca aaccgcggct atgtcgcctt 1620
agccctcggg gtcccacagc ctcagcagcg tcctagcctg cccgctccat gccacggcaa 1680
ggctgcaccg tgttccaggg gtgaaggggg cgatcgggca tgctcctccc catgggtcgc 1740
ccaccatgtc taatggatat cgcactctgt cccagcacct caatgacctg aagaaggaga 1800
acttcagcct caagctgcgc atctacttcc tggaggagcg catgcaacag aagtatgaag 1860
tcagccggga ggacgtctac aagcggaaca ttgagctgaa ggttgaagtg gagagcctga 1920
aacgagagct ccaggacagg aaacagcatc tacataaaac atgggccgat gaggaggatc 1980
tcaacagcca gaatgaagca gagctccggc gccaggttga agaaccgcag caggagacag 2040
aacacgttta tgagctccta gacaacaaca ttcagctgct gcaggaggaa tccaggtttg 2100
caaaggatga agccacacag atggagactc tggtggaggc agagaagggg tgtaatctgg 2160
agctctcaga gaggtggaag gatgctacca agaacaggga agatgcaccg ggagaccagg 2220
tgaagcttga ccaatattct gcggcactgg ctcagaggga caggagaatt gaagagctga 2280
ggcagagctt ggctgcccag gaggggcttg tggaacagct gtctcgagag aaacaacaac 2340
tgttacatct gctggaggag cctgggggca tggaagtgca gcccatgcct aaagggttac 2400
ccacgcaaca aaagccagac ctaaatgaga cccctacaac ccagccatct gtgtctgatt 2460
cccacctggc agaactccag gacaaaatcc agcaaacaga ggtcaccaac aagattcttc 2520
aagagaaact gaatgacatg agctgtgagc tcagatctgc acaggagtcg tctcagaagc 2580
aagatacgac aatccaaagc ctcaaggaaa tgctaaagag cagggaaagt gagactgaag 2640
agctgtacca ggtgattgaa ggtcaaaatg acacaatggc aaagcttccg gaaatgctac 2700
accagagcca gctcggacag ctccagagct cagagggcat tgcccctgct cagcagcaag 2760
tggccctgct tgaccttcag agtgctctgt tctgcagcca gcttgaaatc cagaagctcc 2820
agaggctgtt acgccagaaa gagcgtcagc tggctgacgg caagcggtgc atgcaatttg 2880
tggaggctgc agcacaggag agagagcagc agaaggaagc tgcttggaaa cataaccagg 2940
aattacgaaa agctttgcaa cacctccaag gagaactgca cagtaagagc caacagctcc 3000

```
acgttctgga ggcagaaaaa tataatgaaa ttcgaaccca gggacaaaac attcaacacc  3060
taagtcacag tctgagtcac aaagagcagc taattcagga acttcaggag ctcctacagt  3120
atcgggatac cacagacaaa actctagaca caaatgaggt gtttcttgag aaactacggc  3180
aacgaataca agaccgggca gttgctctag agcgggttat agatgaaaag ttctctgctc  3240
tagaagaaaa ggacaaggaa ctgcggcagc tccggcttgc tgtgagggac cgagaccatg  3300
acttagagag actgcgttgt gtcctgtctg ccaatgaagc taccatgcaa agtatggaga  3360
gtctcctgag ggccagaggc ctggaagtgg agcagttaat tgccacctgc caaaacctcc  3420
agtggttgaa ggaagaattg gaaaccaagt ttggccactg gcagaaggaa caggagagca  3480
tcattcagca gttacagaca tctctgcatg acaggaacaa agaagtagag gatctcagtg  3540
caactttgct ccacaaactt ggacccggcc agagtgaagt agctgaggag ctgtgccagc  3600
gcctgcagcg gaaggaaagg gtgctgcagg accttctgag tgatcggaac aaacaagcca  3660
tggagcacga gatggaggtc cagggactgc tccagtcgat gggcacccgg gaacaggaaa  3720
gacaggctgt tgcagaaaaa atggtacaag ccttcatgga aagaaactcg gaattacagg  3780
ccctgcggca gtatctaggg gggaaggaat taatggcagc atctcaggca ttcatctcta  3840
accaaccagc tggagcgact tctgtaggcc cccaccatgg agagcaaact gaccaaggtt  3900
ctacgcagat gccctctcga gacgacagca cctcgctgac tgccagagag gaggccagca  3960
taccccggtc tacattagga gactcagaca cagttgcacg gctggagaaa gaactgagca  4020
atgccaagga ggagcttgag ctcatggcca aaaaagaaag agaaagccag atagaattgt  4080
ctgccctgca gtccatgatg gctgtgcaag aggaagagct gcaggtgcag gctgctgact  4140
tggagtccct gaccaggaac atacagataa aagaagacct cataaaggac ctgcaaatgc  4200
aactggttga ccctgaagat atgccagcca tggagcgcct gacccaagag gtcttacttc  4260
ttcgggaaaa agttgcttca gtggaacccc agggtcagga agggtcagag aacaggagac  4320
aacagttgct gctgatgtta gaaggactag tggatgaacg gagtcggctc aacgaggccc  4380
tgcaagctga gcggcagctc tacagcagcc tggtcaagtt ccatgcccaa ccagagatct  4440
ctgagagaga ccgaactctg caggtggaac tggaaggggc ccaggtgtta cgcagtcgac  4500
tagaagaagt tcttggaaga agcctggagc gcttaagcag gctggagacc ctggccgcca  4560
ttggaggtgc tactgcaggc gatgagactg aagatacaag cacacagttc acagacagca  4620
ttgaggagga ggctgcacac aacagccacc agcaactcat caaggtgtct ttggagaaaa  4680
gcctgaccac catggagacc cagaacacat gtcttcagcc cccttcccca gtaggagagg  4740
atggtaacag gcatcttcag gaagaaatgc tccacctgag ggctgaaatc caccagccct  4800
tagaagagaa gagaaaagct gaggcagaac tcaaggagct aaaggctcaa attgaggaag  4860
caggattctc ctctgtgtcc cacatcagga acaccatgct gagcctttgc ctttgccttg  4920
agaatgcaga gctgaaagag cagatgggag aagcaatgtc tgatggatgg gaggtggagg  4980
aagacaagga gaagggcgag gtgatggtgg agaccgtggt ggccaaaggg ggtctgagtg  5040
aggacagcct tcaggctgag ttcaggaaag tccaggggag actcaagagt gcctacaaca  5100
tcatcaacct cctcaaagag cagctggtcc tgagaagctc ggaagggaac actaaggaga  5160
tgccagagtt cctcgtgcgc ctggccaggg aggtggacag aatgaacatg ggcttgcctt  5220
cctcggagaa gcatcaacac caagaacagg agaatatgac cgcaaggcct ggccccaggc  5280
cccagagtct caagcttggg acagctctct cagtagacgg ctaccaactg gagaacaagt  5340
cccaggccca agactctgga catcagccag aatttagcct accagggtcc accaaacacc  5400
tgcgctccca gctggctcag tgtagacaac ggtaccaaga tctccaggag aagctgctca  5460
tctcagaagc cactgtgttt gcccaggcaa accagctaga gaagtacaga gccatattaa  5520
gtgaatccct ggtgaagcag gacagcaagc agatccaggt ggaccttcag gacctgggct  5580
atgagacttg tggccgaagt gagaatgaag ctgaacgtga ggagaccacc agccctgagt  5640
gtgaggagca cggtaacctg aagcctgtgg tgctggtgga aggcttgtgc tctgagcaag  5700
ggtacctgga ccctgtcttg gtcagctcac ctgtgaagaa cccttggaga acaagccagg  5760
aagccagaag aatccaggca caaggaactt cagacaacag ctctctcctg aggaaggaca  5820
tccgaaatct gaaagcccag ctaccgaatg cctacaaggt ccttcagaac ctgaggagcc  5880
gggtccggtc cctgtctgcc acaagcgatt actcatcgag tctggagaga ccccgcaagc  5940
tgatagccgt ggcaaccctt gaggggcct caccccacag tgtcactgat gaagacgaag  6000
gcttgttgtc agatggcacc ggggcttttt accctccagg gctccaggcc aaaaagaatc  6060
tagagaatct catccagaga gtatcccagc tggaggccca gctccccaaa actggactag  6120
aagggaagct ggctgaagaa ctgaagtccg cctcgtggcc tggaaaatac gattctttga  6180
ttcaggatca ggcccgaaaa actgtcatat ctgcgtccga aaatacnaaa agggagaagg  6240
atttgttttc ttctcaccca acattcgaaa gatacgtcaa atcttttgaa gacctcctga  6300
ggaacaacga cttgactact tacctgggcc agagcttccg ggaacaactt agttcaaggc  6360
gttcagtgac agacaggctg accagcaaat tcagcacaaa ggatcataag agtgaaaaag  6420
aagaagttgg gcttgagcca ctggccttca ggttcagcag ggaattacag gagaaagaga  6480
aagtgattga agtcctgcag gccaaggtgg atacccggtt tttctcaccc cccagcagcc  6540
atgctgcgtc tgagtcccac cgttgtgcca gcagcacatc tttcctgtcg gatgacatag  6600
aagcctgctc tgacatggac gtagccagcg agtacacaca ctatgaagag aagaagccct  6660
cacccagtaa ctcagcagcc agtgcatctc aggggcttaa gggcgagccc agaagcagct  6720
ccatcagctt gccaactccc cagaaccccc ctaaggaggc cagccaggcc cagccaggct  6780
```

```
ttcactttaa ctccataccc aagccggcta gcctttccca ggcaccaatg cacttcactg    6840
tacccagctt catgcctttc ggcccctctg ggcctccect tcttggttgc tgtgagacac    6900
cagtggtgtc cttggctgag gctcaacaag agctgcagat gctgcagaag cagctgggac    6960
gaagtgttag cattgccect cccacctcca catccacgtt gcttagcaac cacacagaag    7020
ctagctctcc ccgctacagc aaccctgctc agccccactc cccagcaagg ggcaccatag    7080
agctgggcag aatcctggag cctggatacc tgggcagcgg ccagtgggac atgatgaggc    7140
ctcagaaagg gagcatctct ggggagctgt cctcaggctc ctcgatgtac cagcttaact    7200
ccaaacccac aggggccgac ctgttggaag agcatttagg tgagatccgg aacctgcgcc    7260
agcgcctgga ggagtccata tgtgtcaatg acaggctacg ggagcagctg cagcataggc    7320
tcagctccac ggcccgagaa aatggttcca cctctcactt ctacagtcag ggcctggagt    7380
ccatgcctca gctctacaat gagaacagag ccctcaggga agaaaaccaa agcctgcaga    7440
cacggctcag tcatgcttcc aggggacact cccaggaagt ggaccacctg agggaggctc    7500
tgctttcctc aagttcccag ctccaggagc tggagaagga gctggagcag cagaaggctg    7560
agcggcggca gcttctggaa gacttgcagg agaagcagga tgagatcgtg catttccgag    7620
aggagaggct gtccctccag gaaaacaact ccaggctgca gcacaagctg gccctcctgc    7680
aacaacagtg tgaggagaaa cagcagctct ccctgtccct gcagtcagag ctccagatct    7740
acgagtccct ctacgaaaat cctaagaagg gcttgaaagc cttcagccta gattcctgtt    7800
accaagtccc gggtgagttg agctgcctgg tggcagagat tcgagctctg agagtgcagt    7860
tggagcagag cattcaagtg aacaaccgtc tgcggctgca gctggaacag cagatggatc    7920
acggtgctgg caaagccagt ctcagttcct gccctgttaa ccagagcttc tcagccaagg    7980
cggagctggc aaaccagcag ccacccttcc aaggttcagc tgcttccect ccagtccggg    8040
acgttggctt gaattctcca cccgtggtcc tccccagcaa ttcgtgctct gttcctggct    8100
cagactctgc catcatcagt aggacaaaca atggttcgga tgagtctgca gcaacgaaga    8160
ccccctcccaa gatggaggtc gatgctgctg atggcccatt tgccagtgga cacggcagac    8220
acgtcatcgg ccatgtggat gactacgacg ccctacagca gcagattggg gaagggaagc    8280
tgctgatcca aaagatactg tctctcacga ggccagcacg cagcgtccct gcactggacg    8340
cgcagggcac agaggcacca ggtaccaaaa gtgtccatga gcttcggagc agcgccaggg    8400
ctctgaacca cagcctagaa gagtcagctt ccctcctcac catgttctgg agagcagctt    8460
tgccaaactc tcatggttct gtactggtag gcgaagaggg aaacctgatg gagaaagaac    8520
tcctagacct gcgagcccaa gtgtcccaac agcaacagct ccttcagagc actgctgtgc    8580
gtctgaagac ggccaaccag aggaagaaaa gcatggagca gttcatcgtg agccatctga    8640
ccaggaccca tgatgtcttg aagaaagcac ggactaattt agagatgaaa tccttcaggg    8700
ccctgatgtg cactccagcc ttgtgaccct tgccttccag gagccacata aaaggcgaag    8760
ccaggagtcc ttaaaacagc aggaagggtg ggcctgcccg cccctagtac agctgcctgt    8820
ctgctgagga atacctggtc cgactcctcc cctgctggag ctccagggaa gggctcatat    8880
atgtgtccac atgggacagg caggaaggaa agtggcatcc tgacaatgaa tatgattagc    8940
caaggcccac tgggcccatc actaagcaaa actcatgtag actgtgtaga aggccccccg    9000
gcactgcttc tagacagcct cagcagcacg gtgcccacct cgttacagtt ctcacctcaa    9060
gatagccaac tcaggggaac taggacctta ccacccacaa acaggatgtg tggtcccaat    9120
gccaacgctc ctcagacagt tgtaaaagca cacatcattg agtggcagcg tccagccgga    9180
cactgttgga gactaccaaa cccctcactg acccagtctt gggccaggcc agctctgtgg    9240
gccaagtctg gtagtacttt ggtctctacc accacaccag agagagtcta tatagcaaat    9300
gtggtaactt gtaggtgccc tgcacttagc ctagcacctt ctgtttctta cgtgatctca    9360
agttgaacca acttccttaa ctctgctgtc ccctgaatcc taacttccct caggggaatt    9420
ggagattggt ggccacatca tgcctattga atgtttagtg aacagcatat cggtgcctct    9480
taatggcatg ggcaaggcct gctctgtact gaagactgtg tcttcacagt gctcatagga    9540
cgtgggtgtg tgtataaatg tataatatag atttatatat gtcgctatgg ctatgtgttg    9600
aaggccagca taagtgcaga gcgatgggtg agaagacgct aagcagtctt tcttatggct    9660
attaaagcta actgtgtac                                                9679
```

<210> 3
<211> 6981
<212> DNA
<213> rat

<220>
<221> misc_feature
<222> (1)...(6981)
<223> n = A,T,C or G

<400> 3

```
atgtctaatg gatatcgcac tctgtcccag cacctcaatg acctgaagaa ggagaacttc      60
agcctcaagc tgcgcatcta cttcctggag gagcgcatgc aacagaagta tgaagtcagc     120
cgggaggacg tctacaagcg gaacattgag ctgaaggttg aagtggagag cctgaaacga     180
gagctccagg acaggaaaca gcatctacat aaaacatggg ccgatgagga ggatctcaac     240
agccagaatg aagcagagct ccggcgccag gttgaagaac cgcagcagga gacagaacac     300
gtttatgagc tcctagacaa caacattcag ctgctgcagg aggaatccag gtttgcaaag     360
gatgaagcca cacagatgga gactctggtg gaggcagaga aggggtgtaa tctggagctc     420
tcagagaggt ggaaggatgc taccaagaac agggaagatg caccgggaga ccaggtgaag     480
cttgaccaat attctgcggc actggctcag agggacagga gaattgaaga gctgaggcag     540
agcttggctg cccaggaggg gcttgtggaa cagctgtctc gagagaaaca acaactgtta     600
catctgctgg aggagcctgg gggcatggaa gtgcagccca tgcctaaagg gttacccacg     660
caacaaaagc cagacctaaa tgagacccct acaacccagc catctgtgtc tgattcccac     720
ctggcagaac tccaggacaa aatccagcaa acagaggtca ccaacaagat tcttcaagag     780
aaactgaatg acatgagctg tgagctcaga tctgcacagg agtcgtctca gaagcaagat     840
acgacaatcc aaagcctcaa ggaaatgcta aagagcaggg aaagtgagac tgaagagctg     900
taccaggtga ttgaaggtca aaatgacaca atggcaaagc ttccggaaat gctacaccag     960
agccagctcg gacagctcca gagctcagag ggcattgccc ctgctcagca gcaagtggcc    1020
ctgcttgacc ttcagagtgc tctgttctgc agccagcttg aaatccagaa gctccagagg    1080
ctgttacgcc agaaagagcg tcagctggct gacggcaagc ggtgcatgca atttgtggag    1140
gctgcagcac aggagagaga gcagcagaag gaagctgctt ggaaacataa ccaggaatta    1200
cgaaaagctt tgcaacacct ccaaggagaa ctgcacagta agagccaaca gctccacgtt    1260
ctggaggcag aaaaatataa tgaaattcga acccagggac aaaacattca acacctaagt    1320
cacagtctga gtcacaaaga gcagctaatt caggaacttc aggagctcct acagtatcgg    1380
gataccacag acaaaactct agacacaaat gaggtgtttc ttgagaaact acggcaacga    1440
atacaagacc gggcagttgc tctagagcgg gttatagatg aaaagttctc tgctctagaa    1500
gaaaaggaca aggaactgcg gcagctccgg cttgctgtga gggaccgaga ccatgactta    1560
gagagactgc gttgtgtcct gtctgccaat gaagctacca tgcaaagtat ggagagtctc    1620
ctgagggcca gaggcctgga agtggagcag ttaattgcca cctgccaaaa cctccagtgg    1680
ttgaaggaag aattggaaac caagtttggc cactggcaga aggaacagga gagcatcatt    1740
cagcagttac agacatctct gcatgacagg aacaaagaag tagaggatct cagtgcaact    1800
ttgctccaca aacttggacc cggccagagt gaagtagctg aggagctgtg ccagcgcctg    1860
cagcggaagg aaagggtgct gcaggacctt ctgagtgatc ggaacaaaca agccatggag    1920
cacgagatgg aggtccaggg actgctccag tcgatgggca cccgggaaca ggaaagacag    1980
gctgttgcag aaaaaatggt acaagccttc atggaaagaa actcggaatt acaggccctg    2040
cggcagtatc taggggggaa ggaattaatg gcagcatctc aggcattcat ctctaaccaa    2100
ccagctggag cgacttctgt aggcccccac catggagagc aaactgacca aggttctacg    2160
cagatgccct ctcgagacga cagcacctcg ctgactgcca gagaggaggc cagcataccc    2220
cggtctacat taggagactc agacacagtt gcagggctgg agaaagaact gagcaatgcc    2280
aaggaggagc ttgagctcat ggccaaaaaa gaaagagaaa gccagataga attgtctgcc    2340
ctgcagtcca tgatggctgt gcaagaggaa gagctgcagg tgcaggctgc tgacttggag    2400
tccctgacca ggaacataca gataaaagaa gacctcataa aggacctgca aatgcaactg    2460
gttgaccctg aagatatgcc agccatggag cgcctgaccc aagaggtctt acttcttcgg    2520
gaaaaagttg cttcagtgga accccagggt caggaagggt cagagaacag gagacaacag    2580
ttgctgctga tgttagaagg actagtggat gaacggagtc ggctcaacga ggccctgcaa    2640
gctgagcggc agctctacag cagcctggtc aagttccatg cccaaccaga gatctctgag    2700
agagaccgaa ctctgcaggt ggaactggaa ggggcccagg tgttacgcag tcgactagaa    2760
gaagttcttg gaagaagcct ggagcgctta agcaggctgg agaccctggc cgccattgga    2820
ggtgctactg caggcgatga gactgaagat acaagcacac agttcacaga cagcattgag    2880
gaggaggctg cacacaacag ccaccagcaa ctcatcaagg tgtctttgga gaaaagcctg    2940
accaccatgg agacccagaa cacatgtctt cagccccctt ccccagtagg agaggatggt    3000
aacaggcatc ttcaggaaga aatgctccac ctgagggctg aaatccacca gcccttagaa    3060
gagaagagaa aagctgaggc agaactcaag gagctaaagg ctcaaattga ggaagcagga    3120
ttctcctctg tgtcccacat caggaacacc atgctgagcc tttgcctttg ccttgagaat    3180
gcagagctga aagagcagat gggagaagca atgtctgatg gatgggaggt ggaggaagac    3240
aaggagaagg gcgaggtgat ggtggagacc gtggtggcca aaggggtct gagtgaggac     3300
agccttcagg ctgagttcag gaaagtccag gggagactca agagtgccta caacatcatc    3360
aacctcctca aagagcagct ggtcctgaga agctcggaag ggaacactaa ggagatgcca    3420
gagttcctcg tgcgcctggc cagggaggtg gacagaatga acatgggctt gccttcctcg    3480
gagaagcatc aacaccaaga acaggagaat atgaccgcaa ggcctggccc caggccccag    3540
agtctcaagc ttgggacagc tctctcagta gacggctacc aactggagaa caagtcccag    3600
gcccaagact ctggacatca gccagaattt agcctaccag ggtccaccaa acacctgcgc    3660
tcccagctgg ctcagtgtag acaacggtac caagatctcc aggagaagct gctcatctca    3720
gaagccactg tgtttgccca ggcaaaccag ctagagaagt acagagccat attaagtgaa    3780
```

tccctggtga agcaggacag caagcagatc caggtggacc ttcaggacct gggctatgag 3840
acttgtggcc gaagtgagaa tgaagctgaa cgtgaggaga ccaccagccc tgagtgtgag 3900
gagcacggta acctgaagcc tgtggtgctg gtggaaggct tgtgctctga gcaagggtac 3960
ctggaccctg tcttggtcag ctcacctgtg aagaaccctt ggagaacaag ccaggaagcc 4020
agaagaatcc aggcacaagg aacttcagac aacagctctc tcctgaggaa ggacatccga 4080
aatctgaaag cccagctacc gaatgcctac aaggtccttc agaacctgag gagccgggtc 4140
cggtccctgt ctgccacaag cgattactca tcgagtctgg agagaccccg caagctgata 4200
gccgtggcaa cccttgaggg ggcctcaccc cacagtgtca ctgatgaaga cgaaggcttg 4260
ttgtcagatg gcaccggggc tttttaccct ccagggctcc aggccaaaaa gaatctagag 4320
aatctcatcc agagagtatc ccagctggag gcccagctcc ccaaaactgg actagaaggg 4380
aagctggctg aagaactgaa gtccgcctcg tggcctggaa aatacgattc tttgattcag 4440
gatcaggccc gaaaaactgt catatctgcg tccgaaaata cnaaaaggga gaaggatttg 4500
ttttcttctc acccaacatt cgaaagatac gtcaaatctt ttgaagacct cctgaggaac 4560
aacgacttga ctacttacct gggccagagc ttccgggaac aacttagttc aaggcgttca 4620
gtgacagaca ggctgaccag caaattcagc acaaaggatc ataagagtga aaaagaagaa 4680
gttgggcttg agccactggc cttcaggttc agcagggaat tacaggagaa agagaaagtg 4740
attgaagtcc tgcaggccaa ggtggatacc cggtttttct caccccc:ag cagccatgct 4800
gcgtctgagt cccaccgttg tgccagcagc acatctttcc tgtcggatga catagaagcc 4860
tgctctgaca tggacgtagc cagcgagtac acacactatg aagagaagaa gccctcaccc 4920
agtaactcag cagccagtgc atctcagggg cttaagggcg agcccagaag cagctccatc 4980
agcttgccaa ctccccagaa ccccctaag gaggccagcc aggcccagcc aggctttcac 5040
tttaactcca tacccaagcc ggctagcctt tcccaggcac caatgcactt cactgtaccc 5100
agcttcatgc ctttcggccc ctctgggcct ccccttcttg gttgctgtga gacaccagtg 5160
gtgtccttgg ctgaggctca acaagagctg cagatgctgc agaagcagct gggacgaagt 5220
gttagcattg cccctcccac ctccacatcc acgttgctta gcaaccacac agaagctagc 5280
tctccccgct acagcaaccc tgctcagccc cactccccag caaggggcac catagagctg 5340
ggcagaatcc tggagcctgg atacctgggc agcggccagt gggacatgat gaggcctcag 5400
aaagggagca tctctgggga gctgtcctca ggctcctcga tgtaccagct taactccaaa 5460
cccacagggg ccgacctgtt ggaagagcat ttaggtgaga tccggaacct gcgccagcgc 5520
ctggaggagt ccatatgtgt caatgacagg ctacgggagc agctgcagca taggctcagc 5580
tccacggccc gagaaaatgg ttccacctct cacttctaca gtcagggcct ggagtccatg 5640
cctcagctct acaatgagaa cagagccctc agggaagaaa accaaagcct gcagacacgg 5700
ctcagtcatg cttccagggg acactcccag gaagtggacc acctgaggga ggctctgctt 5760
tcctcaagtt cccagctcca ggagctggag aaggagctgg agcagcagaa ggctgagcgg 5820
cggcagcttc tggaagactt gcaggagaag caggatgaga tcgtgcattt ccgagaggag 5880
aggctgtccc tccaggaaaa caactccagg ctgcagcaca agctggccct cctgcaacaa 5940
cagtgtgagg agaaacagca gctctccctg tccctgcagt cagagctcca gatctacgag 6000
tccctctacg aaaatcctaa gaagggcttg aaagccttca gcctagattc ctgttaccaa 6060
gtcccgggtg agttgagctg cctggtggca gagattcgag ctctgagagt gcagttggag 6120
cagagcattc aagtgaacaa ccgtctgcgg ctgcagctgg aacagcagat ggatcacggt 6180
gctggcaaag ccagtctcag ttcctgccct gttaaccaga gcttctcagc caaggcggag 6240
ctggcaaacc agcagccacc cttccaaggt tcagctgctt cccctccagt ccgggacgtt 6300
ggcttgaatt ctccacccgt ggtcctcccc agcaattcgt gctctgttcc tggctcagac 6360
tctgccatca tcagtaggac aaacaatggt tcggatgagt ctgcagcaac gaagacccct 6420
cccaagatgg aggtcgatgc tgctgatggc ccatttgcca gtggacacgg cagacacgtc 6480
atcggccatg tggatgacta cgacgcccta cagcagcaga ttggggaagg gaagctgctg 6540
atccaaaaga tactgtctct cacgaggcca gcacgcagcg tccctgcact ggacgcgcag 6600
ggcacagagg caccaggtac caaaagtgtc catgagcttc ggagcagcgc cagggctctg 6660
aaccacagcc tagaagagtc agcttccctc ctcaccatgt tctggagagc agctttgcca 6720
aactctcatg gttctgtact ggtaggcgaa gagggaaacc tgatggagaa agaactccta 6780
gacctgcgag cccaagtgtc ccaacagcaa cagctccttc agagcactgc tgtgcgtctg 6840
aagacggcca accagaggaa gaaaagcatg gagcagttca tcgtgagcca tctgaccagg 6900
acccatgatg tcttgaagaa agcacggact aatttagaga tgaaatcctt cagggccctg 6960
atgtgcactc cagccttgtg a 6981

<210> 4
<211> 9241
<212> DNA
<213> human

<400> 4

```
ggatccttga gggcactggt gcgactttca ggtgaggtct tagcagatga aagcggctgg      60
ctgtggcccg cgccagtagt gcttttctgct ccgcactcgc cgtgagccag gtgtgcaacc     120
ggatttgggg cgagggtcgc gctggctacc tcgcatgcgc agagccggaa gcccgctgac     180
cggactacag ctcccagaag agccttgtgg aggccgcaga cgcgaagccg ctggcgccat     240
cttgaaatct gatcctccat ccccgaggct ttgcgtctgc gcggccggcc gctgctgctc     300
cgggagccca gtctgctaaa aggggaggac gttgaggacg cggcggctgg cgggagagac     360
agctggggag agacatggca gggtcggagc gcggcctgcg cctctgtcac tcagcatcct     420
cttaggcgtt tccacgcccg ccccctgccc gaggggcggg gctgacggct ctggtacccg     480
gagtcggcgc gcggggcagg ggcgcgcccc tgcagagtgg ggaccccact gggctgtgcc     540
atgctgaccg gagaccaccg aggcgggaga cagagcgcgg cgaagagcca ttgagtggtc     600
acccagtagc cgccgccgcc gccgcctcgg gaagcttgcc acccgctagg agggaagatg     660
aaggagattt gcaggatctg tgcccgagag ctgtgtggaa accagcggcg ctggatcttc     720
cacacggcgt ccaagctcaa tctccaggtt ctgctttcgc acgtcttggg caaggatgtc     780
ccccgcgatg gcaaagccga gttcgcttgc agcaagtgtg ctttcatgct tgatcgaatc     840
tatcgattcg acacagttat tgcccggatt gaagcgcttt ctattgagcg cttgcaaaag     900
ctgctactgg agaaggatcg cctcaagttc tgcattgcca gtatgtatcg gaagaataac     960
gatgactctg gcgcggagat caaggcgggg aatgggacgg ttgacatgtc cgtcttaccc    1020
gatgcgagat actctgcact gctccaggag gacttcgcct attcagggtt tgagtgctgg    1080
gtggagaatg aggatcagat ccaggagcca cacagctgcc atggttcaga aggccctgga    1140
aaccgaccca ggagatgccg tggttgtgcc gctttgcggg ttgctgattc tgactatgaa    1200
gccatttgta aggtacctcg aaaggtggcc agaagtatct cctgcggccc ttctagcagg    1260
tggtcgacca gcatttgcac tgaagaacca gcgttgtctg aggttgggcc acccgactta    1320
gcaagcacaa aggtaccccc agatggagaa agcatggagg aagagacgcc tggttcctct    1380
gtggaatctt tggatgcaag cgtccaggct agccctccac aacagaaaga tgaggagact    1440
gagagaagtg caaaggaact tggaaagtgt gactgttgtt cagatgatca ggctccgcag    1500
catgggtgta atcacaagct ggaattagct cttagcatga ttaaaggtct tgattataag    1560
cccatccaga gcccccgagg gagcaggctt ccgattccag tgaaatccag cctacctgga    1620
gccaagcctg gccctagcat gacagatgga gttagttccg gtttccttaa caggtctttg    1680
aaaccccttt acaagacacc tgtgagttat cccttggagc tttcagacct gcaggagctg    1740
tgggatgatc tctgtgaaga ttatttgccg ctccgggtcc agcccatgac tgaagagttg    1800
ctgaaacaac aaaagctgaa ttcacatgag accactataa ctcagcagtc tgtatctgat    1860
tcccacttgg cagaactcca ggaaaaaatc cagcaaacag aggccaccaa caagattctt    1920
caagagaaac ttaatgaaat gagctatgaa ctaaagtgtg ctcaggagtc gtctcaaaag    1980
caagatggta caattcagaa cctcaaggaa actctgaaaa gcagggaacg tgagactgag    2040
gagttgtacc aggtaattga aggtcaaaat gacacaatgg caaagcttcg agaaatgctg    2100
caccaaagcc agcttggaca acttcacagc tcagagggta cttcttccagc tcagcaacag    2160
gtagctctgc ttgatcttca gagtgcttta ttctgcagcc aacttgaaat acagaagctc    2220
cagagggtgg tacgacagaa agagcgccaa ctggctgatg ccaaacaatg tgtgcaattt    2280
gtagaggctg cagcacacga gagtgaacag cagaaagagg cttcttggaa acataaccag    2340
gaattgcgaa aagccttgca gcagctacaa gaagaattgc agaataagag ccaacagctt    2400
cgtgcctggg aggctgaaaa atacaatgag attcgaaccc aggaacaaaa catccagcac    2460
ctaaaccata gtctgagtca caaggagcag ttgcttcagg aatttcggga gctcctacag    2520
tatcgagata actcagacaa aacccttgaa gcaaatgaaa tgttgcttga gaaacttcgc    2580
cagcgaatac atgataaagc tgttgctctg gagcgggcta tagatgaaaa attctctgct    2640
ctagaagaga aagaaaaaga actgcgccag cttcgtcttg ctgtgagaga gcgagatcat    2700
gacttagaga gactgcgcga tgtcctctcc tccaatgaag ctactatgca aagtatggag    2760
agtctcctga gggccaaagg cctggaagtg gaacagttat ctactacctg tcaaaacctc    2820
cagtggctga aagaagaaat ggaaaccaaa tttagccgtt ggcagaagga acaagagagt    2880
atcattcagc agttacagac gtctcttcat gataggaaca aagaagtgga ggatcttagt    2940
gcaacactgc tctgcaaact tggaccaggg cagagtgaga tagcagagga gctgtgccag    3000
cgtctacagc gaaaggaaag gatgctgcag gaccttctaa gtgatcgaaa taaacaagtg    3060
ctggaacatg aaatggagat tcaaggcctg cttcagtctg tgagcaccag ggagcaggaa    3120
agccaagctg ctgcagagaa gttggtgcaa gccttaatgg aaagaaattc agaattacag    3180
gccctgcgcc aatatttagg agggagagac tccctgatgt cccaagcacc catctctaac    3240
caacaagctg aagttacccc cactggccgt cttggaaaac agactgatca aggttcaatg    3300
cagatacctt ccagagatga tagcacttca ttgactgcca aagaggatgt cagcataccc    3360
agatccacat taggagactt ggacacagtt gcagggctgg aaaaagaact gagtaatgcc    3420
aaagaggaac ttgaactcat ggctaaaaaa gaaagagaaa gtcagatgga actttctgct    3480
ctacagtcca tgatggctgt gcaggaagaa gagctgcagg tgcaggctgc tgatatggag    3540
tctctgacca ggaacataca gattaaagaa gatctcataa aggacctgca aatgcaactg    3600
gttgatcctg aagacatacc agctatggaa cgcctgaccc aggaagtctt acttcttcgg    3660
gaaaaagttg cttcagtaga atcccagggt caagaaattt caggaaaccg aagacaacag    3720
ttgctgctga tgctagaagg actagtagat gaacggagtc ggctcaatga ggccttacaa    3780
```

gcagagagac agctctatag cagtctggtg aagttccatg cccatccaga gagctctgag 3840
agagaccgaa ctctgcaggt ggaactggaa ggggctcagg tgttacgcag tcggctagaa 3900
gaagttcttg gaagaagctt ggagcgctta aacaggctgg agaccctggc cgccattgga 3960
ggtgcagctg caggggatga caccgaagat acaagcactg agttcactga cagtattgag 4020
gaggaggctg cacaccatag tcaccagcaa cttgtcaagg tggctttgga gaaaagtctg 4080
gcaactgtgg agacccagaa cccatctttt tcccctcctt ctccgatggg aggggacagt 4140
aacaggtgtc ttcaggaaga aatgctccac ctgagggctg agttccacca gcacttagaa 4200
gagaagagga aagctgagga ggaactgaag gagctaaagg ctcaaattga ggaagcagga 4260
ttctcctcag tgtcccacat caggaacacc atgctgagcc tttgccttga gaatgcggag 4320
ctgaaagagc agatgggaga agcaatgtct gatggatggg agatcgagga agacaaggag 4380
aagggcgagg tgatggttga gactgtggta accaaagagg gtctgagtga gagtagcctt 4440
caggctgagt tcagaaagct ccagggaaaa ctgaagaatg cccacaatat catcaacctc 4500
ctcaaagaac aacttgtgct gagtagcaag gaagggaata gtaaacttac tccagagctc 4560
cttgtgcatc tgaccagcac cattgaaaga ataaacacag aactggttgg ttcccctggg 4620
aagcaccaac accaagagga ggggaatgtg actgtgaggc ctttccccag accccagagc 4680
cttgaccttg gggctacctt cacagtggat gcccaccaat tggataacca gtcccagcct 4740
cgtgaccctg ggcctcagtc agcgtttagc ctaccagggt ccacccagca cctgcgctcc 4800
cagctgtcac aatgcaaaca acgctatcaa gatctccagg agaagctgct gctatcagaa 4860
gccactgtct ttgctcaggc taacgagctg gagaaataca gagttatgct tacaggtgaa 4920
tccttggtga agcaggacag caagcagatc caggtggacc tccaggacct gggctatgag 4980
acttgtggcc gaagcgagaa tgaggctgaa cgggaggaaa ccaccagtcc tgagtgtgag 5040
gagcacaaca gcctcaagga aatggtcctg atggaggggc tgtgctctga gcagggacgc 5100
cggggctcaa cactggctag ttcctctgag aggaagccct tggagaacca gctagggaag 5160
caggaagagt tccgggtata tggaaagtca gaaaacatct tggtcctacg aaaggacatc 5220
aaagatctga aggcccagct gcagaatgcc aacaaggtca ttcaaaacct caagagccgg 5280
gtccggtccc tctcagttac aagtgattat tcgtctagtc tggaaagacc ccggaagctg 5340
agagctgttg gcaccttgga ggggtcttca cctcatagtg tccctgatga ggatgagggg 5400
tggctgtctg atggcactgg ggctttctac tctccagggc ttcaggccaa aaaggacctg 5460
gagagtctca tccagagagt atcccagctg gaggcccagc tcccaaaaaa tggactagaa 5520
gagaagctgg ctgaggagct gagatcagcc tcgtggcctg ggaaatatga ttccctgatt 5580
caggatcagg cccgggaact gtcttaccta cggcaaaaaa tacgagaagg gagaggtatt 5640
tgttatctta tcacccggca tgcaaaagat acagtaaaat cttttgagga tctcctaagg 5700
agcaatgaca ttgactacta cctgggacag agcttccggg agcaactcgc ccagggaagc 5760
cagctgacag agaggctcac cagcaaactc agcaccaagg atcataaaag tgagaaagat 5820
caagctggac ttgagccact ggccctcagg ctcagcaggg agctgcagga gaaggagaaa 5880
gtgattgaag tcctgcaggc caagctggat gctcggtccc tcacaccctc cagcagccat 5940
gccttgtctg actcccaccg ctctcccagc agcacctctt tcctgtctga tgaactggaa 6000
gcctgctctg acatggacat agtcagcgag tacacacact atgaagagaa gaaagcttct 6060
cccagtcact cagattccat ccatcattcg agtcattctg ctgtgttgtc ttctaaacca 6120
tcatcaacca gtgcatctca gggggctaag gccgaatcca acagcaaccc catcagcttg 6180
ccaactcccc agaatacccc caaggaggcc aaccaggccc attcaggctt tcatttcac 6240
tccataccca agctggctag ccttcctcag gcaccattgc cctcagctcc atccagcttc 6300
ctgcctttca gccccactgg ccctctcctc cttggctgct gtgagacacc agtggtctcc 6360
ttggctgagg ctcagcagga gctacagatg ctgcagaagc agttgggaga aagtgccagc 6420
actgttcctc ctgcttccac agctacattg ctgagcaacg acttggaagc cgactcttcc 6480
tactacctca actctgccca gcctcactct cctccaaggg gcaccataga actgggaaga 6540
atcctagagc ctgggtacct gggcagcagt ggcaagtggg atgtgatgag gcctcagaaa 6600
gggagtgtat ctggggacct atcctcaggc tcctctgtgt accagcttaa ctccaaaccc 6660
acaggggctg acctgctgga agagcatctt ggtgaaatcc ggaacctgcg ccagcgcctg 6720
gaggagtcca tctgcatcaa tgaccgccta cgggagcaac tggaacaccg gctgacctct 6780
actgctcgtg gaaggggatc cacttctaac ttctacagtc agggcctgga gtccatacct 6840
cagctctgca atgagaacag agtcctcagg gaagacaatc gaagacttca ggctcaactg 6900
agtcatgttt ccagagagca ctcccaggaa acagaaagcc tgagggaggc tctgctgtcc 6960
tctcgatccc accttcaaga gctggaaaag gagctggagc accagaaggt ggaaaggcag 7020
cagcttttgg aagacttgag ggagaagcag caagaggtct tgcatttcag ggaggaacgt 7080
ctttccctcc aggaaaacga ctccagtggg ccttgcctct ccctggtcag actgcagcac 7140
aagctggttc tcctgcagca acagtgtgaa gagaaacagc agctctttga gtccctccag 7200
tcagagctac aaatctacga ggcactttat ggcaattcca agaaggggct gaaagcttac 7260
agcctggatg cctgtcacca aatccctttg agcagtgacc tgagccacct ggtggcagag 7320
gtacgagctc tgagagggca gctggagcag agcattcagg ggaacaattg tctgcgactg 7380
cagctgcaac agcagctgga gagcggtgct ggcaaagcca gcctcagccc ctcctccatt 7440
aaccagaact tcccagccag cactgaccct ggaaacaagc agctgctcct ccaagattca 7500
gctgtgtccc ctccagtccg ggatgttggt atgaattccc cagctctggt cttccccagc 7560

```
tctgcttcct ctactcctgg ctcagaaacg cccataatca acagagcaaa tggcttgggt    7620
ttggatactt ctccagtaat gaagacccct cccaagctag agggtgatgc tactgatggc    7680
tcctttgcca ataagcatgg ccgccatgtc attggccaca ttgatgacta cagtgcccta    7740
agacagcaga ttgcggaggg caagctgctg gtcaaaaaga tagtgtctct tgtgagatca    7800
gcgtgcagct tccctggcct tgaagcccaa ggcacagagg tgctaggcag caaaggtatt    7860
catgagcttc ggagcagcac cagtgccctg caccatgccc tagaggagtc ggcttccctc    7920
ctcaccatgt tctggagagc agccctgcca agcacccaca tccctgtgct gcctggcaaa    7980
gtgggagaat caacagaaag ggaacttctg gaactgagaa ccaaagtatc caaacaggag    8040
cggctccttc agagcacaac tgagcatctg aagaacgcca accagcagaa ggagagcatg    8100
gagcagttca tcgtcagcca gctaaccaga acacatgatg ttttaaagaa ggcaaggact    8160
aacttagagg tgaaatccct aagggctctg ccatgtactc cagccttgtg acccttgcct    8220
tccaggaacc atgcaagaag cgcagccacc agaagtcctt aaaacagcag gaaaggtggg    8280
cctgtccccc ttttgtgcag ctacctatct gctgaggagc atctgggcct cattcctcca    8340
agtccacggg agggtccaga agagggagtc agagatgtat cctggtggag ctgggagaaa    8400
ggcagaaagc ctttctgaca gctatggaat acgattagcc aaggtccact tggcccagca    8460
ctaagaaaaa gatgcgtagt ttgcacagaa ggttttgtga tcctgcctct caacagcccc    8520
agcagcttgg gaactagcaa gagcacattt cttgcctcat cagctgtcct gagatggnaa    8580
actcagtgga tataggaccc tgattccgat gaaaggggca cgtggtccca atgctggagc    8640
tcctctggca ggttctaaaa gcacactact gagcagcggt gccctgccgg acactgctgg    8700
cgggggctca gtgagcacta ctcacagatc cacacctgac cctgttgggt cgagtcaggc    8760
tgggccttgg tctgcactgt agcacctgtg ttctttgagt tcacatcatg aatgtggtga    8820
cttcccagat accatctcag gcttaaccta gcacatccta tttcttttct tctatgatat    8880
ccaaattgga ctgacctcac ttcaaagttg ctgtcccatt ttgtcaccct atcttatctc    8940
ggggaaattg cagactgatg gccagaccaa ctctgttgaa attcttgcat agagcaaacc    9000
tgtgctcatt tttaagtggc atgggagagg cccccagcct agtaaagcct agtctgtgtc    9060
ttcacagtgc tggtagaatg tgtttgtgtg tataaatata tgatatagat ttatatatgt    9120
tgctaacgcc atatattgaa ggccaacata actggtggac agggtgggtg acagaaaatg    9180
aaagcctttt tggtgattgt taaagcaaga tgtgtataaa gaaataaata gtttttcttt    9240
c                                                                    9241
```

<210> 5
<211> 2517
<212> PRT
<213> human

<400> 5

```
Met Lys Glu Ile Cys Arg Ile Cys Ala Arg Glu Leu Cys Gly Asn Gln
 1               5                  10                  15
Arg Arg Trp Ile Phe His Thr Ala Ser Lys Leu Asn Leu Gln Val Leu
            20                  25                  30
Leu Ser His Val Leu Gly Lys Asp Val Pro Arg Asp Gly Lys Ala Glu
        35                  40                  45
Phe Ala Cys Ser Lys Cys Ala Phe Met Leu Asp Arg Ile Tyr Arg Phe
    50                  55                  60
Asp Thr Val Ile Ala Arg Ile Glu Ala Leu Ser Ile Glu Arg Leu Gln
65                  70                  75                  80
Lys Leu Leu Leu Glu Lys Asp Arg Leu Lys Phe Cys Ile Ala Ser Met
                85                  90                  95
Tyr Arg Lys Asn Asn Asp Asp Ser Gly Ala Glu Ile Lys Ala Gly Asn
            100                 105                 110
Gly Thr Val Asp Met Ser Val Leu Pro Asp Ala Arg Tyr Ser Ala Leu
        115                 120                 125
Leu Gln Glu Asp Phe Ala Tyr Ser Gly Phe Glu Cys Trp Val Glu Asn
    130                 135                 140
Glu Asp Gln Ile Gln Glu Pro His Ser Cys His Gly Ser Glu Gly Pro
145                 150                 155                 160
Gly Asn Arg Pro Arg Arg Cys Arg Gly Cys Ala Ala Leu Arg Val Ala
                165                 170                 175
Asp Ser Asp Tyr Glu Ala Ile Cys Lys Val Pro Arg Lys Val Ala Arg
            180                 185                 190
Ser Ile Ser Cys Gly Pro Ser Ser Arg Trp Ser Thr Ser Ile Cys Thr
```

```
        195                 200                 205
Glu Glu Pro Ala Leu Ser Glu Val Gly Pro Pro Asp Leu Ala Ser Thr
    210                 215                 220
Lys Val Pro Pro Asp Gly Glu Ser Met Glu Glu Glu Thr Pro Gly Ser
225                 230                 235                 240
Ser Val Glu Ser Leu Asp Ala Ser Val Gln Ala Ser Pro Pro Gln Gln
                245                 250                 255
Lys Asp Glu Glu Thr Glu Arg Ser Ala Lys Glu Leu Gly Lys Cys Asp
            260                 265                 270
Cys Cys Ser Asp Asp Gln Ala Pro Gln His Gly Cys Asn His Lys Leu
        275                 280                 285
Glu Leu Ala Leu Ser Met Ile Lys Gly Leu Asp Tyr Lys Pro Ile Gln
    290                 295                 300
Ser Pro Arg Gly Ser Arg Leu Pro Ile Pro Val Lys Ser Ser Leu Pro
305                 310                 315                 320
Gly Ala Lys Pro Gly Pro Ser Met Thr Asp Gly Val Ser Ser Gly Phe
                325                 330                 335
Leu Asn Arg Ser Leu Lys Pro Leu Tyr Lys Thr Pro Val Ser Tyr Pro
            340                 345                 350
Leu Glu Leu Ser Asp Leu Gln Glu Leu Trp Asp Asp Leu Cys Glu Asp
        355                 360                 365
Tyr Leu Pro Leu Arg Val Gln Pro Met Thr Glu Glu Leu Leu Lys Gln
    370                 375                 380
Gln Lys Leu Asn Ser His Glu Thr Thr Ile Thr Gln Gln Ser Val Ser
385                 390                 395                 400
Asp Ser His Leu Ala Glu Leu Gln Glu Lys Ile Gln Gln Thr Glu Ala
                405                 410                 415
Thr Asn Lys Ile Leu Gln Glu Lys Leu Asn Glu Met Ser Tyr Glu Leu
            420                 425                 430
Lys Cys Ala Gln Glu Ser Ser Gln Lys Gln Asp Gly Thr Ile Gln Asn
        435                 440                 445
Leu Lys Glu Thr Leu Lys Ser Arg Glu Arg Glu Thr Glu Glu Leu Tyr
    450                 455                 460
Gln Val Ile Glu Gly Gln Asn Asp Thr Met Ala Lys Leu Arg Glu Met
465                 470                 475                 480
Leu His Gln Ser Gln Leu Gly Gln Leu His Ser Ser Glu Gly Thr Ser
                485                 490                 495
Pro Ala Gln Gln Gln Val Ala Leu Leu Asp Leu Gln Ser Ala Leu Phe
            500                 505                 510
Cys Ser Gln Leu Glu Ile Gln Lys Leu Gln Arg Val Val Arg Gln Lys
        515                 520                 525
Glu Arg Gln Leu Ala Asp Ala Lys Gln Cys Val Gln Phe Val Glu Ala
    530                 535                 540
Ala Ala His Glu Ser Glu Gln Gln Lys Glu Ala Ser Trp Lys His Asn
545                 550                 555                 560
Gln Glu Leu Arg Lys Ala Leu Gln Gln Leu Gln Glu Glu Leu Gln Asn
                565                 570                 575
Lys Ser Gln Gln Leu Arg Ala Trp Glu Ala Glu Lys Tyr Asn Glu Ile
            580                 585                 590
Arg Thr Gln Glu Gln Asn Ile Gln His Leu Asn His Ser Leu Ser His
        595                 600                 605
Lys Glu Gln Leu Leu Gln Glu Phe Arg Glu Leu Leu Gln Tyr Arg Asp
    610                 615                 620
Asn Ser Asp Lys Thr Leu Glu Ala Asn Glu Met Leu Leu Glu Lys Leu
625                 630                 635                 640
Arg Gln Arg Ile His Asp Lys Ala Val Ala Leu Glu Arg Ala Ile Asp
                645                 650                 655
Glu Lys Phe Ser Ala Leu Glu Glu Lys Glu Lys Glu Leu Arg Gln Leu
            660                 665                 670
Arg Leu Ala Val Arg Glu Arg Asp His Asp Leu Glu Arg Leu Arg Asp
        675                 680                 685
Val Leu Ser Ser Asn Glu Ala Thr Met Gln Ser Met Glu Ser Leu Leu
    690                 695                 700
```

```
Arg Ala Lys Gly Leu Glu Val Glu Gln Leu Ser Thr Thr Cys Gln Asn
705                 710                 715                 720
Leu Gln Trp Leu Lys Glu Glu Met Glu Thr Lys Phe Ser Arg Trp Gln
                725                 730                 735
Lys Glu Gln Glu Ser Ile Ile Gln Gln Leu Gln Thr Ser Leu His Asp
            740                 745                 750
Arg Asn Lys Glu Val Glu Asp Leu Ser Ala Thr Leu Leu Cys Lys Leu
        755                 760                 765
Gly Pro Gly Gln Ser Glu Ile Ala Glu Glu Leu Cys Gln Arg Leu Gln
    770                 775                 780
Arg Lys Glu Arg Met Leu Gln Asp Leu Leu Ser Asp Arg Asn Lys Gln
785                 790                 795                 800
Val Leu Glu His Glu Met Glu Ile Gln Gly Leu Leu Gln Ser Val Ser
                805                 810                 815
Thr Arg Glu Gln Glu Ser Gln Ala Ala Ala Glu Lys Leu Val Gln Ala
            820                 825                 830
Leu Met Glu Arg Asn Ser Glu Leu Gln Ala Leu Arg Gln Tyr Leu Gly
        835                 840                 845
Gly Arg Asp Ser Leu Met Ser Gln Ala Pro Ile Ser Asn Gln Gln Ala
    850                 855                 860
Glu Val Thr Pro Thr Gly Arg Leu Gly Lys Gln Thr Asp Gln Gly Ser
865                 870                 875                 880
Met Gln Ile Pro Ser Arg Asp Asp Ser Thr Ser Leu Thr Ala Lys Glu
                885                 890                 895
Asp Val Ser Ile Pro Arg Ser Thr Leu Gly Asp Leu Asp Thr Val Ala
            900                 905                 910
Gly Leu Glu Lys Glu Leu Ser Asn Ala Lys Glu Glu Leu Glu Leu Met
        915                 920                 925
Ala Lys Lys Glu Arg Glu Ser Gln Met Glu Leu Ser Ala Leu Gln Ser
    930                 935                 940
Met Met Ala Val Gln Glu Glu Glu Leu Gln Val Gln Ala Ala Asp Met
945                 950                 955                 960
Glu Ser Leu Thr Arg Asn Ile Gln Ile Lys Glu Asp Leu Ile Lys Asp
                965                 970                 975
Leu Gln Met Gln Leu Val Asp Pro Glu Asp Ile Pro Ala Met Glu Arg
            980                 985                 990
Leu Thr Gln Glu Val Leu Leu Leu Arg Glu Lys Val Ala Ser Val Glu
        995                 1000                1005
Ser Gln Gly Gln Glu Ile Ser Gly Asn Arg Arg Gln Gln Leu Leu Leu
    1010                1015                1020
Met Leu Glu Gly Leu Val Asp Glu Arg Ser Arg Leu Asn Glu Ala Leu
1025                1030                1035                1040
Gln Ala Glu Arg Gln Leu Tyr Ser Ser Leu Val Lys Phe His Ala His
                1045                1050                1055
Pro Glu Ser Ser Glu Arg Asp Arg Thr Leu Gln Val Glu Leu Glu Gly
            1060                1065                1070
Ala Gln Val Leu Arg Ser Arg Leu Glu Glu Val Leu Gly Arg Ser Leu
        1075                1080                1085
Glu Arg Leu Asn Arg Leu Glu Thr Leu Ala Ala Ile Gly Gly Ala Ala
    1090                1095                1100
Ala Gly Asp Asp Thr Glu Asp Thr Ser Thr Glu Phe Thr Asp Ser Ile
1105                1110                1115                1120
Glu Glu Glu Ala Ala His His Ser His Gln Gln Leu Val Lys Val Ala
                1125                1130                1135
Leu Glu Lys Ser Leu Ala Thr Val Glu Thr Gln Asn Pro Ser Phe Ser
            1140                1145                1150
Pro Pro Ser Pro Met Gly Gly Asp Ser Asn Arg Cys Leu Gln Glu Glu
        1155                1160                1165
Met Leu His Leu Arg Ala Glu Phe His Gln His Leu Glu Glu Lys Arg
    1170                1175                1180
Lys Ala Glu Glu Glu Leu Lys Glu Leu Lys Ala Gln Ile Glu Glu Ala
1185                1190                1195                1200
Gly Phe Ser Ser Val Ser His Ile Arg Asn Thr Met Leu Ser Leu Cys
```

1205 1210 1215
Leu Glu Asn Ala Glu Leu Lys Glu Gln Met Gly Glu Ala Met Ser Asp
1220 1225 1230
Gly Trp Glu Ile Glu Glu Asp Lys Glu Lys Gly Glu Val Met Val Glu
1235 1240 1245
Thr Val Val Thr Lys Glu Gly Leu Ser Glu Ser Ser Leu Gln Ala Glu
1250 1255 1260
Phe Arg Lys Leu Gln Gly Lys Leu Lys Asn Ala His Asn Ile Ile Asn
1265 1270 1275 1280
Leu Leu Lys Glu Gln Leu Val Leu Ser Ser Lys Glu Gly Asn Ser Lys
1285 1290 1295
Leu Thr Pro Glu Leu Leu Val His Leu Thr Ser Thr Ile Glu Arg Ile
1300 1305 1310
Asn Thr Glu Leu Val Gly Ser Pro Gly Lys His Gln His Gln Glu Glu
1315 1320 1325
Gly Asn Val Thr Val Arg Pro Phe Pro Arg Pro Gln Ser Leu Asp Leu
1330 1335 1340
Gly Ala Thr Phe Thr Val Asp Ala His Gln Leu Asp Asn Gln Ser Gln
1345 1350 1355 1360
Pro Arg Asp Pro Gly Pro Gln Ser Ala Phe Ser Leu Pro Gly Ser Thr
1365 1370 1375
Gln His Leu Arg Ser Gln Leu Ser Gln Cys Lys Gln Arg Tyr Gln Asp
1380 1385 1390
Leu Gln Glu Lys Leu Leu Leu Ser Glu Ala Thr Val Phe Ala Gln Ala
1395 1400 1405
Asn Glu Leu Glu Lys Tyr Arg Val Met Leu Thr Gly Glu Ser Leu Val
1410 1415 1420
Lys Gln Asp Ser Lys Gln Ile Gln Val Asp Leu Gln Asp Leu Gly Tyr
1425 1430 1435 1440
Glu Thr Cys Gly Arg Ser Glu Asn Glu Ala Glu Arg Glu Glu Thr Thr
1445 1450 1455
Ser Pro Glu Cys Glu Glu His Asn Ser Leu Lys Glu Met Val Leu Met
1460 1465 1470
Glu Gly Leu Cys Ser Glu Gln Gly Arg Arg Gly Ser Thr Leu Ala Ser
1475 1480 1485
Ser Ser Glu Arg Lys Pro Leu Glu Asn Gln Leu Gly Lys Gln Glu Glu
1490 1495 1500
Phe Arg Val Tyr Gly Lys Ser Glu Asn Ile Leu Val Leu Arg Lys Asp
1505 1510 1515 1520
Ile Lys Asp Leu Lys Ala Gln Leu Gln Asn Ala Asn Lys Val Ile Gln
1525 1530 1535
Asn Leu Lys Ser Arg Val Arg Ser Leu Ser Val Thr Ser Asp Tyr Ser
1540 1545 1550
Ser Ser Leu Glu Arg Pro Arg Lys Leu Arg Ala Val Gly Thr Leu Glu
1555 1560 1565
Gly Ser Ser Pro His Ser Val Pro Asp Glu Asp Glu Gly Trp Leu Ser
1570 1575 1580
Asp Gly Thr Gly Ala Phe Tyr Ser Pro Gly Leu Gln Ala Lys Lys Asp
1585 1590 1595 1600
Leu Glu Ser Leu Ile Gln Arg Val Ser Gln Leu Glu Ala Gln Leu Pro
1605 1610 1615
Lys Asn Gly Leu Glu Glu Lys Leu Ala Glu Glu Leu Arg Ser Ala Ser
1620 1625 1630
Trp Pro Gly Lys Tyr Asp Ser Leu Ile Gln Asp Gln Ala Arg Glu Leu
1635 1640 1645
Ser Tyr Leu Arg Gln Lys Ile Arg Glu Gly Arg Gly Ile Cys Tyr Leu
1650 1655 1660
Ile Thr Arg His Ala Lys Asp Thr Val Lys Ser Phe Glu Asp Leu Leu
1665 1670 1675 1680
Arg Ser Asn Asp Ile Asp Tyr Tyr Leu Gly Gln Ser Phe Arg Glu Gln
1685 1690 1695
Leu Ala Gln Gly Ser Gln Leu Thr Glu Arg Leu Thr Ser Lys Leu Ser
1700 1705 1710

Thr Lys Asp His Lys Ser Glu Lys Asp Gln Ala Gly Leu Glu Pro Leu
1715 1720 1725
Ala Leu Arg Leu Ser Arg Glu Leu Gln Glu Lys Glu Lys Val Ile Glu
1730 1735 1740
Val Leu Gln Ala Lys Leu Asp Ala Arg Ser Leu Thr Pro Ser Ser Ser
1745 1750 1755 1760
His Ala Leu Ser Asp Ser His Arg Ser Pro Ser Ser Thr Ser Phe Leu
1765 1770 1775
Ser Asp Glu Leu Glu Ala Cys Ser Asp Met Asp Ile Val Ser Glu Tyr
1780 1785 1790
Thr His Tyr Glu Glu Lys Lys Ala Ser Pro Ser His Ser Asp Ser Ile
1795 1800 1805
His His Ser Ser His Ser Ala Val Leu Ser Ser Lys Pro Ser Ser Thr
1810 1815 1820
Ser Ala Ser Gln Gly Ala Lys Ala Glu Ser Asn Ser Asn Pro Ile Ser
1825 1830 1835 1840
Leu Pro Thr Pro Gln Asn Thr Pro Lys Glu Ala Asn Gln Ala His Ser
1845 1850 1855
Gly Phe His Phe His Ser Ile Pro Lys Leu Ala Ser Leu Pro Gln Ala
1860 1865 1870
Pro Leu Pro Ser Ala Pro Ser Ser Phe Leu Pro Phe Ser Pro Thr Gly
1875 1880 1885
Pro Leu Leu Leu Gly Cys Cys Glu Thr Pro Val Val Ser Leu Ala Glu
1890 1895 1900
Ala Gln Gln Glu Leu Gln Met Leu Gln Lys Gln Leu Gly Glu Ser Ala
1905 1910 1915 1920
Ser Thr Val Pro Pro Ala Ser Thr Ala Thr Leu Leu Ser Asn Asp Leu
1925 1930 1935
Glu Ala Asp Ser Ser Tyr Tyr Leu Asn Ser Ala Gln Pro His Ser Pro
1940 1945 1950
Pro Arg Gly Thr Ile Glu Leu Gly Arg Ile Leu Glu Pro Gly Tyr Leu
1955 1960 1965
Gly Ser Ser Gly Lys Trp Asp Val Met Arg Pro Gln Lys Gly Ser Val
1970 1975 1980
Ser Gly Asp Leu Ser Ser Gly Ser Ser Val Tyr Gln Leu Asn Ser Lys
1985 1990 1995 2000
Pro Thr Gly Ala Asp Leu Leu Glu Glu His Leu Gly Glu Ile Arg Asn
2005 2010 2015
Leu Arg Gln Arg Leu Glu Glu Ser Ile Cys Ile Asn Asp Arg Leu Arg
2020 2025 2030
Glu Gln Leu Glu His Arg Leu Thr Ser Thr Ala Arg Gly Arg Gly Ser
2035 2040 2045
Thr Ser Asn Phe Tyr Ser Gln Gly Leu Glu Ser Ile Pro Gln Leu Cys
2050 2055 2060
Asn Glu Asn Arg Val Leu Arg Glu Asp Asn Arg Arg Leu Gln Ala Gln
2065 2070 2075 2080
Leu Ser His Val Ser Arg Glu His Ser Gln Glu Thr Glu Ser Leu Arg
2085 2090 2095
Glu Ala Leu Leu Ser Ser Arg Ser His Leu Gln Glu Leu Glu Lys Glu
2100 2105 2110
Leu Glu His Gln Lys Val Glu Arg Gln Gln Leu Leu Glu Asp Leu Arg
2115 2120 2125
Glu Lys Gln Gln Glu Val Leu His Phe Arg Glu Glu Arg Leu Ser Leu
2130 2135 2140
Gln Glu Asn Asp Ser Ser Gly Pro Cys Leu Ser Leu Val Arg Leu Gln
2145 2150 2155 2160
His Lys Leu Val Leu Leu Gln Gln Gln Cys Glu Glu Lys Gln Gln Leu
2165 2170 2175
Phe Glu Ser Leu Gln Ser Glu Leu Gln Ile Tyr Glu Ala Leu Tyr Gly
2180 2185 2190
Asn Ser Lys Lys Gly Leu Lys Ala Tyr Ser Leu Asp Ala Cys His Gln
2195 2200 2205
Ile Pro Leu Ser Ser Asp Leu Ser His Leu Val Ala Glu Val Arg Ala

```
    2210                2215                2220
Leu Arg Gly Gln Leu Glu Gln Ser Ile Gln Gly Asn Asn Cys Leu Arg
2225                2230                2235                2240
Leu Gln Leu Gln Gln Gln Leu Glu Ser Gly Ala Gly Lys Ala Ser Leu
                2245                2250                2255
Ser Pro Ser Ser Ile Asn Gln Asn Phe Pro Ala Ser Thr Asp Pro Gly
            2260                2265                2270
Asn Lys Gln Leu Leu Leu Gln Asp Ser Ala Val Ser Pro Pro Val Arg
        2275                2280                2285
Asp Val Gly Met Asn Ser Pro Ala Leu Val Phe Pro Ser Ser Ala Ser
    2290                2295                2300
Ser Thr Pro Gly Ser Glu Thr Pro Ile Ile Asn Arg Ala Asn Gly Leu
2305                2310                2315                2320
Gly Leu Asp Thr Ser Pro Val Met Lys Thr Pro Pro Lys Leu Glu Gly
                2325                2330                2335
Asp Ala Thr Asp Gly Ser Phe Ala Asn Lys His Gly Arg His Val Ile
            2340                2345                2350
Gly His Ile Asp Asp Tyr Ser Ala Leu Arg Gln Gln Ile Ala Glu Gly
        2355                2360                2365
Lys Leu Leu Val Lys Lys Ile Val Ser Leu Val Arg Ser Ala Cys Ser
    2370                2375                2380
Phe Pro Gly Leu Glu Ala Gln Gly Thr Glu Val Leu Gly Ser Lys Gly
2385                2390                2395                2400
Ile His Glu Leu Arg Ser Ser Thr Ser Ala Leu His His Ala Leu Glu
                2405                2410                2415
Glu Ser Ala Ser Leu Leu Thr Met Phe Trp Arg Ala Ala Leu Pro Ser
            2420                2425                2430
Thr His Ile Pro Val Leu Pro Gly Lys Val Gly Glu Ser Thr Glu Arg
        2435                2440                2445
Glu Leu Leu Glu Leu Arg Thr Lys Val Ser Lys Gln Glu Arg Leu Leu
    2450                2455                2460
Gln Ser Thr Thr Glu His Leu Lys Asn Ala Asn Gln Gln Lys Glu Ser
2465                2470                2475                2480
Met Glu Gln Phe Ile Val Ser Gln Leu Thr Arg Thr His Asp Val Leu
                2485                2490                2495
Lys Lys Ala Arg Thr Asn Leu Glu Val Lys Ser Leu Arg Ala Leu Pro
            2500                2505                2510
Cys Thr Pro Ala Leu
        2515
```

<210> 6
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primers

<400> 6
cggaattcga ggaggcctac cagaaac 27

<210> 7
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primers

<400> 7
tgagtcgact acgtgtcaag gcaacaatgg tc 32

<210> 8

<211> 1683
<212> PRT
<213> rat

<400> 8

```
Met Met Ala Gln Phe Pro Thr Ala Met Asn Gly Gly Pro Asn Met Trp
 1               5                  10                  15
Ala Ile Thr Ser Glu Glu Arg Thr Lys His Asp Lys Gln Phe Asp Asn
            20                  25                  30
Leu Lys Pro Ser Gly Gly Tyr Ile Thr Gly Asp Gln Ala Arg Thr Phe
        35                  40                  45
Phe Leu Gln Ser Gly Leu Pro Ala Pro Val Leu Ala Glu Ile Trp Ala
    50                  55                  60
Leu Ser Asp Leu Asn Lys Asp Gly Lys Met Asp Gln Gln Glu Phe Ser
65                  70                  75                  80
Ile Ala Met Lys Leu Ile Lys Leu Lys Leu Gln Gly Gln Gln Leu Pro
                85                  90                  95
Val Val Leu Pro Pro Ile Met Lys Gln Pro Pro Met Phe Ser Pro Leu
            100                 105                 110
Ile Ser Ala Arg Phe Gly Met Gly Ser Met Pro Asn Leu Ser Ile His
        115                 120                 125
Gln Pro Leu Pro Pro Val Ala Pro Ile Thr Ala Pro Leu Ser Ser Ala
    130                 135                 140
Thr Ser Gly Thr Ser Ile Pro Pro Leu Met Met Pro Ala Pro Leu Val
145                 150                 155                 160
Pro Ser Val Ser Thr Ser Ser Leu Pro Asn Gly Thr Ala Ser Leu Ile
                165                 170                 175
Gln Pro Leu Ser Ile Pro Tyr Ser Ser Ser Thr Leu Pro His Ala Ser
            180                 185                 190
Ser Tyr Ser Leu Met Met Gly Gly Phe Gly Gly Ala Ser Ile Gln Lys
        195                 200                 205
Ala Gln Ser Leu Ile Asp Leu Gly Ser Ser Ser Ser Thr Ser Ser Thr
    210                 215                 220
Ala Ser Leu Ser Gly Asn Ser Pro Lys Thr Gly Thr Ser Glu Trp Ala
225                 230                 235                 240
Val Pro Gln Pro Ser Arg Leu Lys Tyr Arg Gln Lys Phe Asn Ser Leu
                245                 250                 255
Asp Lys Ser Met Ser Gly Tyr Leu Ser Gly Phe Gln Ala Arg Asn Ala
            260                 265                 270
Leu Leu Gln Ser Asn Leu Ser Gln Thr Gln Leu Ala Thr Ile Trp Thr
        275                 280                 285
Leu Ala Asp Ile Asp Gly Asp Gly Gln Leu Lys Ala Glu Glu Phe Ile
    290                 295                 300
Leu Ala Met His Leu Thr Asp Met Ala Lys Ala Gly Gln Pro Leu Pro
305                 310                 315                 320
Leu Thr Leu Pro Pro Glu Leu Val Pro Pro Ser Phe Arg Gly Gly Lys
                325                 330                 335
Gln Ile Asp Ser Ile Asn Gly Thr Leu Pro Ser Tyr Gln Lys Thr Gln
            340                 345                 350
Glu Glu Glu Pro Gln Lys Lys Leu Pro Val Thr Phe Glu Asp Lys Arg
        355                 360                 365
Lys Ala Asn Tyr Glu Arg Gly Asn Met Glu Leu Glu Lys Arg Arg Gln
    370                 375                 380
Val Leu Met Glu Gln Gln Gln Arg Glu Ala Glu Arg Lys Ala Gln Lys
385                 390                 395                 400
Glu Lys Glu Glu Trp Glu Arg Lys Gln Arg Glu Leu Gln Glu Gln Glu
                405                 410                 415
Trp Lys Lys Gln Leu Glu Leu Glu Lys Arg Leu Glu Lys Gln Arg Glu
            420                 425                 430
Leu Glu Arg Gln Arg Glu Glu Glu Arg Arg Lys Glu Ile Glu Arg Arg
        435                 440                 445
Glu Ala Ala Lys Gln Glu Leu Glu Arg Gln Arg Arg Leu Glu Trp Glu
    450                 455                 460
```

```
Arg Ile Arg Arg Gln Glu Leu Leu Asn Gln Lys Asn Arg Glu Gln Glu
465                 470                 475                 480
Glu Ile Val Arg Leu Asn Ser Lys Lys Lys Ser Leu His Leu Glu Leu
                485                 490                 495
Glu Ala Val Asn Gly Lys His Gln Gln Ile Ser Gly Arg Leu Gln Asp
            500                 505                 510
Val Arg Ile Arg Lys Gln Thr Gln Lys Thr Glu Leu Glu Val Leu Asp
        515                 520                 525
Lys Gln Cys Asp Leu Glu Ile Met Glu Ile Lys Gln Leu Gln Gln Glu
    530                 535                 540
Leu Gln Glu Tyr Gln Asn Lys Leu Ile Tyr Leu Val Pro Glu Lys Gln
545                 550                 555                 560
Leu Leu Asn Glu Arg Ile Lys Asn Met Gln Leu Ser Asn Thr Pro Asp
                565                 570                 575
Ser Gly Ile Ser Leu Leu His Lys Lys Ser Ser Glu Lys Glu Glu Leu
            580                 585                 590
Cys Gln Arg Leu Lys Glu Gln Leu Asp Ala Leu Glu Lys Glu Thr Ala
        595                 600                 605
Ser Lys Leu Ser Glu Met Asp Ser Phe Asn Asn Gln Leu Lys Cys Gly
    610                 615                 620
Asn Met Asp Asp Ser Val Leu Gln Cys Leu Leu Ser Leu Leu Ser Cys
625                 630                 635                 640
Leu Asn Asn Leu Phe Leu Leu Leu Lys Glu Leu Arg Glu Ser Tyr Asn
                645                 650                 655
Thr Gln Gln Leu Ala Leu Glu Gln Leu His Lys Ile Lys Arg Asp Lys
            660                 665                 670
Leu Lys Glu Leu Glu Arg Lys Arg Leu Glu Gln Ile Gln Lys Lys Lys
        675                 680                 685
Leu Glu Asp Glu Ala Ala Arg Lys Ala Lys Gln Gly Lys Glu Asn Leu
    690                 695                 700
Trp Lys Glu Ser Ile Arg Lys Glu Glu Glu Glu Lys Gln Lys Arg Leu
705                 710                 715                 720
Gln Glu Glu Lys Ser Gln Asp Arg Thr Gln Glu Glu Glu Arg Lys Thr
                725                 730                 735
Glu Ala Lys Gln Ser Glu Thr Ala Arg Ala Leu Val Asn Tyr Arg Ala
            740                 745                 750
Leu Tyr Pro Phe Glu Ala Arg Asn His Asp Glu Met Ser Phe Asn Ser
        755                 760                 765
Gly Asp Ile Ile Gln Val Asp Glu Lys Thr Val Gly Glu Pro Gly Trp
    770                 775                 780
Leu Tyr Gly Ser Phe Gln Gly Lys Phe Gly Trp Phe Pro Cys Asn Tyr
785                 790                 795                 800
Val Glu Lys Met Leu Ser Ser Asp Lys Thr Pro Ser Pro Lys Lys Ala
                805                 810                 815
Leu Leu Pro Pro Ala Val Ser Leu Ser Ala Thr Ser Ala Ala Pro Gln
            820                 825                 830
Pro Leu Cys Ser Asn Gln Pro Ala Pro Val Thr Asp Tyr Gln Asn Val
        835                 840                 845
Ser Phe Ser Asn Leu Asn Val Asn Thr Thr Trp Gln Gln Lys Ser Ala
    850                 855                 860
Phe Thr Arg Thr Val Ser Pro Gly Ser Val Ser Pro Ile His Gly Gln
865                 870                 875                 880
Gly Gln Ala Val Glu Asn Leu Lys Ala Gln Ala Leu Cys Ser Trp Thr
                885                 890                 895
Ala Lys Lys Glu Asn His Leu Asn Phe Ser Lys His Asp Val Ile Thr
            900                 905                 910
Val Leu Glu Gln Gln Glu Asn Trp Trp Phe Gly Glu Val His Gly Gly
        915                 920                 925
Arg Gly Trp Phe Pro Lys Ser Tyr Val Lys Ile Ile Pro Gly Ser Glu
    930                 935                 940
Val Lys Arg Gly Glu Pro Glu Ala Leu Tyr Ala Ala Val Asn Lys Lys
945                 950                 955                 960
Pro Thr Ser Thr Ala Tyr Pro Val Gly Glu Glu Tyr Ile Ala Leu Tyr
```

965 970 975
Ser Tyr Ser Ser Val Glu Pro Gly Asp Leu Thr Phe Thr Glu Gly Glu
980 985 990
Glu Leu Leu Val Thr Gln Lys Asp Gly Glu Trp Trp Thr Gly Ser Ile
995 1000 1005
Gly Glu Arg Thr Gly Ile Phe Pro Ser Asn Tyr Val Arg Pro Lys Asp
1010 1015 1020
Gln Glu Asn Val Gly Asn Ala Ser Lys Ser Gly Ala Ser Asn Lys Lys
1025 1030 1035 1040
Pro Glu Ile Ala Gln Val Thr Ser Ala Tyr Ala Ala Ser Gly Ala Glu
1045 1050 1055
Gln Leu Ser Leu Ala Pro Gly Gln Leu Ile Leu Ile Leu Lys Lys Asn
1060 1065 1070
Ser Ser Gly Trp Trp Gln Gly Glu Leu Gln Ala Arg Gly Lys Lys Arg
1075 1080 1085
Gln Lys Gly Trp Phe Pro Ala Ser His Val Lys Leu Leu Gly Pro Ser
1090 1095 1100
Ala Glu Arg Thr Thr Pro Ala Phe His Ala Val Cys Gln Val Ile Ala
1105 1110 1115 1120
Met Tyr Asp Tyr Ile Ala Asn Asn Glu Asp Glu Leu Asn Phe Ser Lys
1125 1130 1135
Gly Gln Leu Ile Asn Val Met Asn Lys Asp Asp Pro Asp Trp Trp Gln
1140 1145 1150
Gly Glu Ile Asn Gly Val Thr Gly Leu Phe Pro Ser Asn Tyr Val Lys
1155 1160 1165
Met Thr Thr Asp Ser Asp Pro Ser Gln Gln Trp Cys Ala Asp Leu Gln
1170 1175 1180
Ala Leu Asp Thr Met Gln Pro Met Glu Arg Lys Arg Gln Gly Tyr Ile
1185 1190 1195 1200
His Glu Leu Ile Glu Thr Glu Glu Arg Tyr Met Asp Asp Leu Gln Leu
1205 1210 1215
Val Ile Glu Val Phe Gln Lys Arg Met Ala Glu Ser Gly Phe Leu Thr
1220 1225 1230
Glu Ala Glu Met Ala Leu Ile Phe Val Asn Trp Lys Glu Leu Ile Met
1235 1240 1245
Ser Asn Thr Lys Leu Leu Lys Ala Leu Arg Val Arg Lys Lys Thr Gly
1250 1255 1260
Gly Glu Lys Met Pro Val Glu Met Met Gly Asp Ile Leu Ala Ala Glu
1265 1270 1275 1280
Leu Ser His Met Gln Ala Tyr Ile Arg Phe Cys Ser Cys Gln Leu Asn
1285 1290 1295
Gly Ala Ala Leu Leu Gln Gln Lys Thr Asp Glu Asp Ala Asp Phe Lys
1300 1305 1310
Glu Phe Leu Lys Lys Leu Ala Ser Asp Pro Arg Cys Lys Gly Met Pro
1315 1320 1325
Leu Ser Ser Phe Leu Leu Lys Pro Met Gln Arg Ile Thr Arg Tyr Pro
1330 1335 1340
Leu Leu Ile Arg Ser Ile Leu Glu Asn Thr Pro Gln Asn His Val Asp
1345 1350 1355 1360
His Ser Ser Leu Lys Leu Ala Leu Glu Arg Ala Glu Glu Leu Cys Ser
1365 1370 1375
Gln Val Asn Glu Gly Val Arg Glu Lys Glu Asn Ser Asp Arg Leu Glu
1380 1385 1390
Trp Ile Gln Ala His Val Gln Cys Glu Gly Leu Ala Glu Gln Leu Ile
1395 1400 1405
Phe Asn Ser Leu Thr Asn Cys Leu Gly Pro Arg Lys Leu Leu Tyr Ser
1410 1415 1420
Gly Lys Leu Tyr Lys Thr Lys Ser Asn Lys Glu Leu His Gly Phe Leu
1425 1430 1435 1440
Phe Asn Asp Phe Leu Leu Leu Thr Tyr Leu Val Arg Gln Phe Ala Ala
1445 1450 1455
Ser Ser Gly Phe Glu Lys Leu Phe Ser Ser Lys Ser Ser Ala Gln Phe
1460 1465 1470

```
Lys Met Tyr Lys Thr Pro Ile Phe Leu Asn Glu Val Leu Val Lys Leu
        1475                1480                1485
Pro Thr Asp Pro Ser Ser Asp Glu Pro Val Phe His Ile Ser His Ile
    1490                1495                1500
Asp Arg Val Tyr Thr Leu Arg Thr Asp Asn Ile Asn Glu Arg Thr Ala
1505                1510                1515                1520
Trp Val Gln Lys Ile Lys Ala Ala Ser Glu Gln Tyr Ile Asp Thr Glu
                1525                1530                1535
Lys Lys Lys Arg Glu Lys Ala Tyr Gln Ala Arg Ser Gln Lys Thr Ser
            1540                1545                1550
Gly Ile Gly Arg Leu Met Val His Val Ile Glu Ala Thr Glu Leu Lys
        1555                1560                1565
Ala Cys Lys Pro Asn Gly Lys Ser Asn Pro Tyr Cys Glu Ile Ser Met
    1570                1575                1580
Gly Ser Gln Ser Tyr Thr Thr Arg Thr Leu Gln Asp Thr Leu Asn Pro
1585                1590                1595                1600
Lys Trp Asn Phe Asn Cys Gln Phe Phe Ile Lys Asp Leu Tyr Gln Asp
                1605                1610                1615
Val Leu Cys Leu Thr Met Phe Asp Arg Asp Gln Phe Ser Pro Asp Asp
            1620                1625                1630
Phe Leu Gly Arg Thr Glu Val Pro Val Ala Lys Ile Arg Thr Glu Gln
        1635                1640                1645
Glu Ser Lys Gly Pro Thr Thr Arg Arg Leu Leu Leu His Glu Val Pro
    1650                1655                1660
Thr Gly Glu Val Trp Val Arg Phe Asp Leu Gln Leu Phe Glu Gln Lys
1665                1670                1675                1680
Thr Leu Leu
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/26860

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :C07H 21/00
US CL : 536/23.2
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

BIOSIS CA CAPLUS EMBASE MEDLINE GENBANK SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KALCHMAN, M.A. HIP1, a human homologue of S.cerevisiae Sla2p, interacts with membrane-associated huntingtin in the brain. Nature Genetics. May 1997, Vol. 16, No. 1 pages 44-53, entire document.	1-3, 9-12.
X	Database GenBank Accession No. 075042. SEKI, N. et al. 'Characterization of cDNA clones in size-fractionated cDNA libraries from human brain'., 01 November 1998.	1-3
X	Database GenBank Accession No. 075065. SEKI, N. et al. 'Characterization of cDNA clones in size-fractionated cDNA libraries from human brain'. 01 November 1998.	1-3
X	Database GenBank Accession No. AA987244. NCI-CGAP 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index'. 27, July 1998.	3

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
19 JANUARY 2000	10 FEB 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer MANJUNATH RAO Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/26860

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GenBank Accession No. AA664799. NCI-CGAP 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gen Index'. 13, February 1998.	3
X	Database GenBank Accession No. AB007923. OHARA, O. 'Homo sapiens mRNA for KIAAA0454 protein, partial cds.' 13, August 1998.	
X	GenBank Accession No. AB007946. O'HARA et al. 'Homo sapiens male brain cDNA to mRNA, clone lib:pBluescriptII SK plus clone:HH0492'. 13 August 1998.	3
X	Database GenBank Accession No. AA671390. MARRA et al. 'The WashU-HHMI Mouse EST Project'. 25 November 1997	3
X	Database GenBank Accession No. AA110441. MARRA, M. et al. 'The WashU-HHMI Mouse EST Project'. 03 February 1997.	3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/26860

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3 and 9-12

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/26860

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-3 and 9-12, drawn to polynucleotides encoding PDE-binding proteins.
Group II, claims 4-8, drawn to PDE-binding proteins.
Group III, claims 13-15, drawn to a monoclonal antibody.
Group IV, claims 16-19, drawn to a method of determining the agent that modulates PDE activity.
Group V, claim 20, drawn to a method of modulating PDE activity.

The inventions listed as Groups I-V do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The polynucleotides encoding PDE-interacting proteins are known in the prior art and does not contribute over the prior art (Kalchman et al. Nature Genetics, May 1997, Vol. 16(1):44-53).

Group I is a product; this shares the special technical feature of DNA molecules which groups II-V do not share.

Group II is a product; this shares the special technical feature of a protein which groups I and III-V do not share.

Group III is a product; this shares the special technical feature of an antibody which groups I,II, IV-V do not share.

Groups IV and V are processes; this shares the special technical feature of uncharacterized chemical compounds which groups I-III do not share.